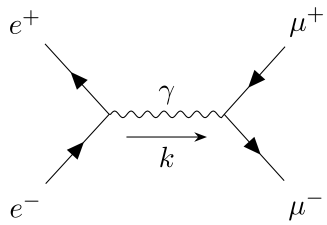
Quantum Field Theory

Lecture Notes

Instructor: Prof. Ritesh Kumar Singh



Sagnik Seth

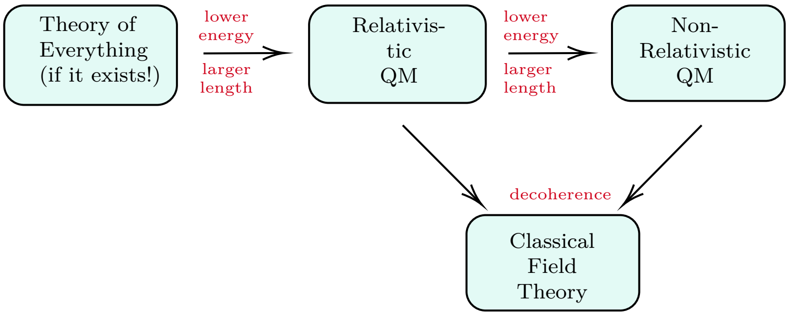
Dept of Physical Sciences
IISER Kolkata

Contents

Lecture 01: Intro and Quantum SHO

People do 'weird' stuffs for earning,
You can do the same for learning!

Many physical theories have been proposed and it is imperative to see how each theory is linked with the previous in some way.



From the proposed theory of everything, we can transform to a *relativistic* quantum field theory which will be the topic of discussion for most part. When subjected to *decoherence*, a relativistic QFT collapse to a classical QFT as noise hides the 'quantumness' from consideration. Hence we will start our discussion from classical fields and then move to quantum fields.

Since the foundational aspect of QFT (and many other topics in Physics) is the Harmonic Oscillator, let us discuss the quantum Harmonic oscillator for introduction. For that, note that the classical Hamiltonian for the harmonic oscillator is given by:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

We do not like the things like m and ω which prevent us from seeing things clearly 😅. Hence, we invoke the holy action of making things dimensionless. For that, we define:

$$X = \sqrt{\frac{m\omega}{\hbar}}x \qquad \longrightarrow \qquad x^2 = \frac{\hbar}{m\omega}X^2$$
$$P = \frac{1}{\sqrt{m\omega\hbar}}p \qquad \longrightarrow \qquad p^2 = m\omega\hbar P^2$$

Substituting these in the Hamiltonian, we have:

$$H = \frac{\hbar\omega}{2}(X^2 + P^2)$$

Note that, we are still within the classical domain. Now, let us elevate x and p to operators and we define:

$$[x, p] = i\hbar\, \mathbb{1} \implies [X, P] = \sqrt{\frac{\cancel{m\omega}}{\hbar}}\frac{i\hbar\, \mathbb{1}}{\sqrt{\cancel{m\omega\hbar}}} = i$$

Introducing the commutator bracket brings us to the quantum world. Now, we invoke our very own ladder operators:

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}) \qquad : \text{annihilation operator}$$
$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P}) \qquad : \text{creation operator}$$

Then we will have [1]:

$$\left[a, a^\dagger\right] = \frac{1}{2}\left[X + iP, X - iP\right] = \frac{-i}{2}([X, P] - [P, X]) = -i \times i = 1$$

[1]henceforth, forsaking the hat symbol and identity operator $\mathbb{1}$, since they cause nothing but trouble, when the context is clear

Also, note that:

$$a^\dagger a = \frac{1}{2}(X^2 + P^2 + i\underbrace{(XP - PX)}_{i}) = \frac{1}{2}(X^2 + P^2) - \frac{1}{2} \implies H = \frac{\hbar\omega}{2}(a^\dagger a + \frac{1}{2})$$

Let us consider the *complete set of commuting observables* (CSCO) for this problem. Evidently, the set $\{H\}$ itself satisfies the condition since the eigenvalues are all non-degenerate (hence, we can label each state with only one index). To understand why, let us consider the action of the the annihilation operator on a (normalised) state $|\psi\rangle$. For that we note the following:

$$\left[a^\dagger a, a\right] = -a \implies [H, a] = -\hbar\omega\ a$$
$$\left[a^\dagger a, a\right] = a^\dagger \implies \left[H, a^\dagger\right] = \hbar\omega\ a^\dagger$$

Now, we have:

$$Ha\,|\psi\rangle - aH\,|\psi\rangle = [H, a]\,|\psi\rangle = -\hbar\omega a\,|\psi\rangle \implies H(a\,|\psi\rangle) = (E - \hbar\omega)(a\,|\psi\rangle)$$

Thus, if $|\psi\rangle$ has an energy eigenvalue E, then $a\,|\psi\rangle$ will have an energy eigenvalue $E - \hbar\omega$. Thus, starting from any eenrgy state, we can change to another state with energy reduced by one unit of $\hbar\omega$, using the annihilation operator. Similarly, we will have:

$$H(a^\dagger\,|\psi\rangle) = (E + \hbar\omega)(a^\dagger\,|\psi\rangle)$$

Let us denote the states $a^\dagger\,|\psi\rangle$ and $a\,|\psi\rangle$ by $|\psi_+\rangle$ and $|\psi_-\rangle$ respectively. Then, we will have

$$\langle\psi_-|\psi_-\rangle = \langle\psi|\,a^\dagger a\,|\psi\rangle = \langle\psi|\left(\frac{H}{\hbar\omega} - \frac{1}{2}\right)|\psi\rangle$$

Now, since $|\psi_-\rangle$ is a valid vector of the Hilbert space, its norm must be non-negative and finite. Hence, we obtain the condition:

$$0 \leq \frac{E}{\hbar\omega} - \frac{1}{2} < \infty \implies \frac{\hbar\omega}{2} \leq E$$

Hence, we get a lower bound on the energy eigenvalue, that is, there exists a state $|\psi_{\text{min}}\rangle$ such that $H\,|\psi_{\text{min}}\rangle = E_{min}\,|\psi_{\text{min}}\rangle$ where $E_{min} = \frac{\hbar\omega}{2}$. Then, starting from this state, if we apply the creation operator, we will get successively increasing energies (and hence the system is non-degenerate). We thus obtain:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad n = 0, 1, \ldots$$

Lecture 02: Dimensional Deep Dive

For the quantum harmonic oscillator, we had seen that how making things dimensionless eases calculations a bit. We consider three fundamental constants, each occupying their own special place in their field of action.

$$c \equiv [LT^{-1}] \qquad : \text{relativity}$$
$$\hbar \equiv [L^2MT^{-1}] \qquad : \text{quantum}$$
$$G \equiv [L^3M^{-1}T^{-2}] \qquad : \text{gravity}$$

Lengths are tractable for us, since we can see the ‘length’, same goes for mass, atleast we can ‘feel’ it. However, time is an *enigma*. *We can’t hold two ends of time at the same time* unlike holding two ends of a rod to measure its length. The absolute truth is: *Time passes!*

Note that, in physics, we are mostly concerned with equations like $E = mc^2$ and $E = \hbar\omega$. To that extend, we define the natural units:

$$c = 1$$
$$\hbar = 1$$

We want these quantitites to be numerically equal to 1 and dimensionless. Note that, we had also done this kind of things before. When writing Newton's law, we had said that

$$F \propto m, F \propto a \implies F \propto ma \implies F = kma \qquad \text{for some } k$$

Now, we chose unit and dimension of force in a way such that $k = 1$ (dimensionless and unit value) which gave us the celebrated law.

Note that:

- Making c dimensionless:
$$[LT^{-1}] \equiv 0 \implies [L] = [T]$$
- Making $\hbar$ dimensionless:
$$[L^2MT^{-1}] = [L^2ML^{-1}] = [LM] \equiv 0 \implies [L] = [M^{-1}]$$

Hence, we see that length and time are equivalent while mass and length have inverse relation. In this natural units, we have that energy is equivalent to mass and any other unit can be represented in terms of mass. Thus, by our convention, we choose mass or energy as the only important dimension.

Note that
$$\hbar c \equiv Jm = 1 \qquad \text{(in natural units)}$$

From this, we can say heuristically that increasing length scale is decreases the energy (mass) scale. We mainly use the length scale in context of *de Broglie wavelength.*

This is perhaps not the only way to define natural units. In cosmology, G plays a more important role and hence it is better to set $G = 1$, leaving aside $\hbar$. Using this *natural units*, we find:

$$[L] = [M]$$
$$[L] = [T]$$

Here, we see that length and mass scale are directly related. Well, in this regard, we treat the length scale to be that of the Schwarzschild radius of a blackhole [1] which intuitively grows with mass. Since the above two natural units are widely distinct, there is as such no problem, however, in the unfortunate case where we have to consider both de Broglie wavelength and the Schwarzschild radius (that is, in the infamous domain of quantum gravity 😨), one needs to be very careful.

Theorem 1 ($\pi-$Theorem):
Let $q_1, \dots, q_n$ be n variables which are physically relevant to a problem and which are related by an expression, that is,

$$F(q_1, \dots, q_n) = 0 \iff q_i = \widetilde{F}(q_1, \dots \hat{q}_i, \dots q_n)$$

If k is the number of fundamental dimensions required to describe the n variables, then we can group these in $(n-k)$ groups of dimensionless variables $\Pi_1, \dots, \Pi_{n-k}$ such that for some f, we have:

$$f(\Pi_1, \dots, \Pi_{n-k}) = 0 \iff \Pi_i = \widetilde{f}(\Pi_1, \dots \hat{\Pi}_i, \dots \Pi_{n-k})$$

.

[1] This crap is the radius of an object such that if the body is squeezed to a radius lesser than the Schwarzschild radius, the gravitational attraction between the constituents of the body causes its irreversible collapse, turning it to a black hole ⚫

The theorem seems a bit vague (and pointless too). Let us take a physical example. Consider a spherical ball in a viscous fluid. The variable in the problem are:

$$\begin{aligned}
\text{Drag force:} \quad & q_1 \to & F \quad & [MLT^{-2}] \\
\text{Sphere diameter:} \quad & q_2 \to & d \quad & [L] \\
\text{Fluid density:} \quad & q_3 \to & \rho \quad & [ML^{-3}] \\
\text{Fluid velocity:} \quad & q_4 \to & v \quad & [LT^{-1}] \\
\text{Fluid viscosity:} \quad & q_5 \to & \eta \quad & [ML^{-1}T^{-1}]
\end{aligned}$$

So, there are 5 such parameters and only three units viz. M, L, T are needed to describe them. Hence we will have two Π groups. It is a good thing to choose the repeating variables (variables which will be in both groups) that relate to mass, geometry and the kinematics of the problem. Also, note that since the Π groups are dimensionless, we can take the non-repeating variable's power to be 1. Hence, in this problem we choose them to be ρ, d, v. Thus, we will have:

$$\Pi_1 = \rho^{a_1} d^{a_2} v^{a_3} F \equiv [ML^{-3}]^{a_1}[L]^{a_2}[LT^{-1}]^{a_3}[MLT^{-2}] = [M^{a_1+1}L^{-3a_1+a_2+a_3+1}T^{-a_3-2}]$$
$$\Pi_2 = \rho^{b_1} d^{b_2} v^{b_3} \eta \equiv [ML^{-3}]^{b_1}[L]^{b_2}[LT^{-1}]^{b_3}[ML^{-1}T^{-1}] = [M^{b_1+1}L^{-3b_1+b_2+b_3-1}T^{-b_3-1}]$$

Hence we obtain two sets of equations:

$$\begin{aligned}
a_1 + 1 = 0 &\implies a_1 = -1 \\
-a_3 - 2 = 0 &\implies a_3 = -2 \\
-3a_1 + a_2 + a_3 + 1 = 0 &\implies 3 + a_2 - 2 + 1 = 0 \implies a_2 = -2 \\
b_1 &= -1 \\
b_3 &= -1 \\
-3b_1 + b_2 + b_3 - 1 = 0 &\implies 3 + b_2 - 2 = 0 \implies b_2 = -1
\end{aligned}$$

Then we obtain the Π groups as:

$$\Pi_1 = \frac{F}{\rho d^2 v^2} \qquad \Pi_2 = \frac{\eta}{\rho d v} = \frac{1}{\frac{\rho d v}{\eta}}$$

We identity $\frac{\rho d v}{\eta}$ to be the *Reynold's number* $\mathcal{R}$. Then we can say, for some ϕ,

$$\frac{F}{\rho d^2 v^2} = \phi(\mathcal{R})$$

To some extraterrestrial being, the physical laws will (hopefully) be valid, however, they might not understand the human-made units (like metres and seconds) in which we measure these quantitites.

We need something natural, based on Nature and hence we used the natural units. For this purpose, we will use things like $c, \hbar, G, k_b \ldots$ and we set all of them to 1. Doing this will lead to change in all scales. For that, we define the Planck units, made of fundamental constants of Nature:

- **Planck mass**: $E_p = \sqrt{\frac{\hbar c}{G}}$
- **Planck length**: $l_p = \sqrt{\frac{\hbar G}{c^3}}$
- **Planck time**: $t_p = \sqrt{\frac{\hbar c}{G^5}}$

Let us now analyse the physical regimes in which the fundamental constants become important. For that, we will use a tuple $(G, \frac{1}{c}, \hbar)$ (note that all the elements of the tuple are written in terms of very small quantitites of the SI scale).

$$
\begin{aligned}
(0,0,0) &\rightarrow \text{Classical mechanics} \\
(G,0,0) &\rightarrow \text{Newtonian gravity} \\
\left(0,\tfrac{1}{c},0\right) &\rightarrow \text{Special relativity} \\
(0,0,\hbar) &\rightarrow \text{Basic quantum mechanics} \\
\left(G,\tfrac{1}{c},0\right) &\rightarrow \text{General relativity} \\
\left(0,\tfrac{1}{c},\hbar\right) &\rightarrow \text{QFT and relativistic QM} \\
(G,0,\hbar) &\rightarrow \text{Non-relativistic gravity} \\
\left(G,\tfrac{1}{c},\hbar\right) &\rightarrow \text{Quantum gravity}
\end{aligned}
$$

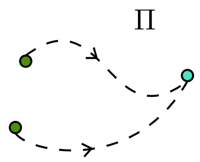
Lecture 03: From Classical to Quantum

Inspecting the transition from a classical to quantum description is necessary for understanding QFT, which is just a *sophisticated mechanics.*

We know that the position $\mathbf{q} \equiv \{q_1, \ldots, q_n\}$ and momentum $\mathbf{p} \equiv \{p_1, \ldots, p_n\}$ uniquely define the state of a *classical* particle.

$$(\mathbf{q}, \mathbf{p}) \in \Pi \text{ (}2n\text{ dimensional phase space)} \qquad q_i, p_i \in \mathbb{R}$$

The Hamiltonian $H(\mathbf{q}, \mathbf{p})$ is just a way of performing time evolution on the system. Using the Hamiltonian, we can 'fibrate' the phase space, that is, starting from one point, we can go to some other point. Also, note that, changing the parameters in the Hamiltonian (eg. ω in SHO Hamiltonian), we change the fibration patterns(eg. *Lissajou's* figures).



In the quantum world, the phase space changes to the Hilbert space $\mathcal{H}$ while a point in the phase space (q,p) changes to a vector $|\psi\rangle$ in the Hilbert space.

Also, these real variables become Hermitian (self-adjoint) operators [1]and the classical Poisson bracket now transforms to the commutator.

$$
\begin{aligned}
\text{CM:} &\qquad \{q,p\} = 1 \\
\text{QM:} &\qquad [\hat{q},\hat{p}] = i
\end{aligned}
$$

3.1. Time Evolution

From classical Hamilton's equations of motion, we have:

$$
\begin{aligned}
\dot{q}_i &= \{q_i, H\} = \frac{\partial H}{\partial p_i} \\
\dot{p}_i &= \{p_i, H\} = -\frac{\partial H}{\partial q_i}
\end{aligned}
$$

[1] A self-adjoint operator $\mathcal{O}$ is such that $\mathcal{O} = \mathcal{O}^\dagger$ in all respect, that is, $\mathcal{O}$ and $\mathcal{O}^\dagger$ have the same domain and action. In general, $\mathbb{D}(\mathcal{O}) \subseteq \mathbb{D}(\mathcal{O}^\dagger)$ where $\mathbb{D}(\cdot)$ represents the domain of some operator. If the domains are not equal but action is same on a restricted domain (mainly occurs in infinite-dimensional spaces), then those operators are not self-adjoint but called *symmetric/Hermitian* (in some places Hermiticity also requires a symmetric operator to be bounded).

If $\mathbf{z} = \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix}$ is a $2n$ dimensional vector, then $\dot{\mathbf{z}} = \{\mathbf{z}, H\}$. Now, moving to quantum mechanics, we know the celebrated Schrödinger equation, which specifies the time evolution of a state vector $|\psi(t)\rangle \in \mathcal{H}$:

$$i\frac{\partial |\psi(t)\rangle}{\partial t} = H |\psi(t)\rangle$$

Thus, in both classical and quantum case, the time derivative of a quantity is equal to some action of the Hamiltonian on that quantity (classical → Poisson bracket, quantum → Multiplication with Hamiltonian)[1].

The Hamiltonian in the position basis becomes:

$$H = \frac{-\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{x})$$

Thus, the Schrödinger equation now acts on the wavefunction $\psi(\mathbf{x}, t) = \langle \mathbf{x}|\psi\rangle$:

$$i\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\frac{1}{2m}\boldsymbol{\nabla}^2\psi(\mathbf{x}, t) + V(\mathbf{x})\psi(x, t)$$

And the energy and momentum operators become $E \to i\frac{\partial}{\partial t}$ and $p \to -i\boldsymbol{\nabla}$.

3.1.1. Free Particle

The dispersion relation for a non-relativistic free particle is:

$$E = \frac{\mathbf{p}^2}{2m}$$

The typical solution can be written as plane-waves, $e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} = e^{i(\mathbf{p}\cdot\mathbf{x}-Et)} \equiv \exp(-ip^\mu x_\mu)$ where we have used the repeated summation notation and taken the metric $\eta_{\mu\nu} = (+, -, -, -)$.

For a relativistic free particle, the dispersion relation becomes:

$$E^2 = m^2 + p^2$$

Instead of the wavefunction $\psi(\mathbf{x}, t)$, we will use $\phi(\mathbf{x}, t)$, since the resulting equation actually acts on a scalar-field and ϕ is common notation for a scalar-field. Substituting the energy and momentum operators here, we get:

$$\begin{aligned}
i^2\frac{\partial^2\phi(\mathbf{x}, t)}{\partial t^2} &= (-i)^2\boldsymbol{\nabla}^2\phi(\mathbf{x}, t) + m^2\phi(\mathbf{x}, t) \\
\implies \frac{\partial^2\phi(\mathbf{x}, t)}{\partial t^2} - \boldsymbol{\nabla}^2\phi(\mathbf{x}, t) &= -m^2\phi(\mathbf{x}, t) \\
\implies \left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + m^2\right)\phi(\mathbf{x}, t) = 0
\end{aligned}$$

The above equation is called the *Klein-Gordon equation.* In covariant notation, we can compactly write it as:

$$\left(\partial_\mu\partial^\mu + m^2\right)\phi(x^\nu) = 0$$

The KG equation is Lorentz invariant, that is, under a Lorentz transformation, $x^\mu \to x'^\nu$, $\phi'(x'^\nu) = \phi(x^\mu)$

[1] A better analogy would be to use density matrices which gives the von Neumann equation which uses commutator bracket

3.1.2. Continuity Equation

Taking the complex cojungate of the Schrödinger's equation and then after some algebraic manipulation, we obtain:

$$\frac{\partial(\psi^*\psi)}{\partial t} = \frac{i}{2m}\boldsymbol{\nabla}\cdot[\psi^*(\boldsymbol{\nabla}\psi) - \psi(\boldsymbol{\nabla}\psi^*)]$$

Here, we identify:

$$\rho := \psi^*\psi \qquad\qquad \mathbf{J} = \frac{-i}{2m}[\psi^*(\boldsymbol{\nabla}\psi) - \psi(\boldsymbol{\nabla}\psi^*)]$$

which yields the well-known form of the *continuity equation*:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{J} = 0 \qquad \longrightarrow \qquad \partial_\mu J^\mu = 0$$

where $J^\mu = (\rho, \mathbf{J})$ is the four-current. Note that in this case, ρ is a positive-definite quantity and can indeed have the interpretation of probability. Also, note that if ψ somehow becomes real-valued, then $\mathbf{J} = 0$ which implies that ρ is constant in time (though it can change in space).

Doing the same thing to Klein-Gordon equation yields:

$$\frac{\partial}{\partial t}\left(\phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t}\right) - \boldsymbol{\nabla}\cdot(\phi^*\boldsymbol{\nabla}\phi - \phi\boldsymbol{\nabla}\phi^*) = 0$$

From this equation, we can identify:

$$\rho := \left(\phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t}\right) \qquad\qquad \mathbf{J} := -(\phi^*\boldsymbol{\nabla}\phi - \phi\boldsymbol{\nabla}\phi^*)$$

Thus here, the 4-current becomes $J^\mu = (\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*)$ Note the apparent problems with this identification:

- It is not apriori obvious that ρ is positive-definite and hence has problem with probability interpretation.
- The equation is second-order in time and hence ρ seems to have a term evolving forward and one term evolving backward in time.
- The dispersion relation does not have a single solution; the solutions are $\pm E$

The thing is, KG equation treats time and space on equal footing (unlike Schrödinger equation where time was in first order and space was in second order). Hence we have to consider all possibilities of moving back and forth in both space and time.

Lecture 04: Classical Fields

Let us start with a familiar setting, the spring-mass system in one-dimensions. So, each mass is connected with isotropic springs of natural length a and spring constant k, and at some instance, i^{th} spring is displaced from its equilibrium position by an amount η_i as shown below:

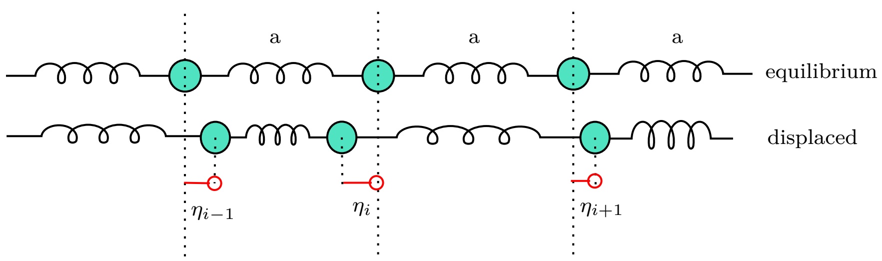


Figure 1: A spring-mass system; each mass is displaced by some amount from its equilibrium position

Then the kinetic and potential energies can be written as:

$$T = \frac{1}{2}\sum_i \dot{\eta}_i^2 \qquad V = \frac{1}{2}\sum_i k(\eta_{i+} - \eta_i)^2$$

From this, the Lagrangian of the system is obtained as:

$$L = T - V = \frac{1}{2}\sum_i a\left[({}^m/_a)\dot{\eta}_i^2 - ka\left(\frac{\eta_{i+1} - \eta_i}{a}\right)^2\right] = a\sum_i L_i$$

The resulting ELEOM are:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\eta}_i}\right) = \frac{\partial L}{\partial \eta} \implies ({}^m/_a)\ddot{\eta}_i - ka\left(\frac{\eta_{i+1} - \eta_i}{a^2}\right) + ka\left(\frac{\eta_i - \eta_{i-1}}{a^2}\right) = 0$$

In the *continuum* limit, that is, $a \to 0$ (so, the system kind of becomes like a rod), we have:

$$\frac{m}{a} \longrightarrow \mu \qquad \text{: mass per unit length of the rod}$$

$$ka \longrightarrow Y \qquad \text{: Young's modulus of the rod}$$

$$i \longrightarrow x \qquad \text{: position label}$$

$$\eta_i \longrightarrow \eta(x) \qquad \text{: field}$$

$$\frac{\eta_{i+1} - \eta_i}{a} \longrightarrow \frac{d\eta}{dx} \qquad \text{: derivative}$$

$$\sum_i \longrightarrow \int dx \qquad \text{: integral}$$

Hence in the continuum limit, the ELEOM becomes:

$$\lim_{a\to 0}\left(\mu\frac{\mathrm{d}^2\eta}{\mathrm{d}t^2} - \frac{Y}{a}\frac{\mathrm{d}\eta}{\mathrm{d}x}\bigg|_x + \frac{Y}{a}\frac{\mathrm{d}\eta}{\mathrm{d}x}\bigg|_{x-a}\right) = 0 \implies \mu\frac{\mathrm{d}^2\eta}{\mathrm{d}t^2} - Y\frac{\mathrm{d}^2\eta}{\mathrm{d}x^2} = 0$$

Defining the velocity as $v := \sqrt{\frac{Y}{\mu}}$, we obtain the wave-equation:

$$\frac{1}{v^2}\frac{\mathrm{d}^2\eta}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\eta}{\mathrm{d}x^2} = 0$$

Note that, in this case, position becomes a mere label like in the discrete case, i just labelled the index of the i^{th} mass. Also, technically, η can also depend on time, so it is better to write it as $\eta(x,t)$.

Position and time, both are kind of labels and thus are independent. Hence, total and partial derivatives coincide in this case. Note that in continuum, the Lagrangian itself becomes an integral of the form:

$$L = \frac{1}{2}\int\left[\mu\dot{\eta}^2 - Y\left(\frac{\mathrm{d}\eta}{\mathrm{d}x}\right)^2\right]dx$$

The bracketted term is the Lagrangian density, denoted by $\mathcal{L}$ and hence, the Lagrangian is the space integral of the Lagrangian density:

$$L = \int \mathcal{L}\, dx$$

Henceforth, we allow ourselves to be *sloppy* enough and refer to Lagrangian density as simply the Lagrangian when context is clear.

4.1. Our actions tell a lot!

Let us now consider a general setting. So, the Lagrangian density can be a function of the field, the spatial and temporal derivatives and space and time themselves. In compact notation, we write:

$$\mathcal{L} \equiv \mathcal{L}(\phi, \partial_\mu \phi, x^\mu)$$

Remember, that we define the classical action S as:

$$S = \int_{t_1}^{t_2} dt\ L = \int_{t_1}^{t_2} dt \int d^3\mathbf{x}\ \mathcal{L}(\phi, \partial_\mu\phi, x^\mu) = \int_{\mathcal{R}\subset\mathbb{R}^4} d^4\mathbf{x}\ \mathcal{L}(\phi, \partial_\mu\phi, x^\mu)$$

The integration can indeed be over a region $\mathcal{R}$ which is a subset of $\mathbb{R}^4$ in general. Mostly, we take $\mathcal{R} = [t_1, t_2] \times \mathbb{R}^3$.

We now want to use Hamilton's principle, which states that, under fixed boundary conditions, the system takes a path in configuration space for which the action is stationary, that is,

$$\delta S = 0$$

Note that the action is a functional which takes a field ϕ as input and returns a real number[1]. Let us consider a variation in the field itself: $\phi \longrightarrow\ \phi' = \phi + \delta\phi$ such that $\delta\phi$ vanishes at the boundary of the integration domain. Then we have the following:

$$\delta S = S[\phi'] - S[\phi] = \int d^4\mathbf{x}\ \big(\mathcal{L}(\phi', \partial_\mu\phi', x^\mu) - \mathcal{L}(\phi, \partial_\mu\phi, x^\mu)\big)$$

Now, we try to find what the above quantity is:

$$\begin{aligned}\mathcal{L}(\phi', \partial_\mu\phi', x^\mu) &= \mathcal{L}(\phi + \delta\phi, \partial_\mu(\phi+\delta\phi), x^\mu) \\ &= \mathcal{L}(\phi, \partial_\mu\phi, x^\mu) + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi)\end{aligned}$$

Also note that $\delta(\partial_\mu\phi) = \partial_\mu(\delta\phi)$. Hence, from the above expression, we obtain that:

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

Using *integration by parts* on the second term, we obtain:

$$\int d^4\mathbf{x}\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) = -\int d^4\mathbf{x}\ \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right)\Bigg|_{\text{boundary}}$$

The variation in the field vanishes at the boundary by our assumption. Then we obtain, for the variation of the action:

$$\delta S = \int d^4\mathbf{x}\ \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right)\delta\phi = 0$$

Since the variation in the field was arbitrary, we obtain the Euler-Lagrange's equations for the field as:

$$\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) = 0$$

In general, we want the action to be *Lorentz* invariant, in order to maintain the universality of physical laws. Since the volume element of Minkowski space $d^4\mathbf{x}$ is invariant under Lorentz transformations, this imply that $\mathcal{L}$ should also be constructed of Lorentz-invariant quantities.

[1] For functionals, a common notation for the input is to use square brackets, that is, if A is a linear functional taking function f as input, $A[f] \in \mathbb{R}$

The field ϕ (and all its powers) is a Lorentz *scalar* and is hence invariant. The derivative $\partial_\mu \phi$ is not invariant, however, $\partial^\mu \partial_\mu \phi$ is. Bilinears like $(\partial^\mu \phi)(\partial_\mu \phi)$ are also invariant. Hence, a simple Lagrangian should be constructed of these quantities. Let us consider the Lagrangian:

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

Then, from the Lagrange's equations,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \phi} &= -m^2\phi \\
\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} &= \frac{1}{2}\eta^{\mu\nu}(\partial_\nu \phi + \partial_\mu \phi \delta^\mu{}_\nu) \\
&= \frac{1}{2}\eta^{\mu\nu}(2 \times \partial_\nu \phi) \\
&= \eta^{\mu\nu}\partial_\nu \phi \\
&= \partial^\mu \phi
\end{aligned}$$

Then we finally obtain:

$$-m^2\phi - \partial_\mu \partial^\mu \phi = 0 \implies (\partial_\mu \partial^\mu + m^2)\phi = 0$$

This is the KG equation that we had seen earlier. Thus, the above Lagrangian models the KG equation.

When we have multiple fields $\{\phi_i\}$, then for each field, we will obtain:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) = 0$$

NOTE: In terms of the total Lagrangian, the ELEOM for the field is:

$$\frac{\delta L}{\delta \phi} - \partial_t \left(\frac{\delta L}{\delta \dot{\phi}} \right) = 0$$

Lecture 05 & 06: Looking into Hamiltonians

Previously, we saw a ton of crap about the Lagrangians and how the equations of motions are obtained from the action being stationary. Now, we will look into the *Hamiltonian* formalism which will allow us to quantise the field.

In basic classical mechanics, we had seen that given a Lagrangian $L(q_i, \dot{q}_i, t)$, we obtained the Hamiltonian by the Legendre transformation as:

$$H(q_i, p_i, t) = p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

where $p_i := \frac{\partial L}{\partial \dot{q}_i}$ was defined to be the generalised momenta, conjugate to the generalised coordinate q_i. Similarly, for classical fields, we can define:

$$\pi^i(x^\mu) := \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi_i)} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_i}$$

This is the *canonical momentum* corresponding to the field. From the ELEOM, we obtain $\dot{\pi}^i = \frac{\delta L}{\delta \phi}$ Using this, the *Hamiltonian density* can be defined as:

$$\mathcal{H}(\phi^a, \pi_a, \partial_i \phi_a, x^\mu) := \pi^a \dot{\phi}_a - \mathcal{L}(\phi_a, \partial_\mu \phi_a, x^\mu)$$

In the final expression on the LHS, we have to replace $\dot{\phi}_a$ in terms of π^a to get the Hamiltonian density (Replacing this will get rid of all the time derivative and hence, in the Hamiltonian, we used ∂_i, the spatial derivatives only). Using this then, we get the total Hamiltonian as:

$$H[\pi^a, \phi_a] = \int d^3\mathbf{x} \; \mathcal{H}(\phi_a, \pi_a, \partial_i \phi_a, x^\mu)$$

6.1. Functional Derivatives

Since action, Lagrangian, etc. are functionals, an apt way to describe their variation is through the *functional derivative*. These come under the domain of functional analysis (which is Math ☠️), however, a short introduction might be necessary.

Let us take a functional $\mathcal{F}[f]$ and consider its variation $\delta\mathcal{F}$ when $f \to f + \delta f$, where $\delta f = \epsilon\eta$. Here ϵ is a small parameter and η is a generic function, often called the *test function*. Now,

$$\mathcal{F}[f+\epsilon\eta] = \mathcal{F}[f] + \left.\frac{\mathrm{d}\mathcal{F}[f+\epsilon\eta]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \epsilon + \mathcal{O}(\epsilon^2)$$

We define the functional derivative as the first-order coefficient of the expansion, that is,

$$\frac{\delta\mathcal{F}}{\delta f} = \lim_{\epsilon\to 0} \frac{\mathcal{F}[f+\epsilon\eta] - \mathcal{F}[f]}{\epsilon}$$

Functional derivatives seem to follow similar rules as in ordinary differential calculus (viz. linearity, Liebnitz rule, chain rule) and hence, can be treated to be operationally same as the ordinary derivative. However, exceptions are always there.

The differential (variation) of a functional is given as:

$$\delta\mathcal{F}[\phi] := \int \frac{\delta\mathcal{F}[\phi]}{\delta\phi}\delta\phi$$

Now let us try to relate functional derivative with normal derivative. For that, let us suppose that the space is discretised into cells, each of volume ΔV^i and each cell is assigned the *average value* of ϕ within the volume, that is,

$$\phi_i(t) = \frac{1}{\Delta V^i}\int\limits_{\Delta V^i} d^3\mathbf{x}\ \phi(\mathbf{x},t)$$

Then, L now depends on the discrete components, ϕ_i and $\dot{\phi}_i$ and the variation can be written as in terms of partial derivatives (since ϕ_i are now discrete):

$$\delta L = \sum_i \left(\frac{\partial L}{\partial\phi_i}\delta\phi_i + \frac{\partial L}{\partial\dot{\phi}_i}\delta\dot{\phi}_i\right) = \sum_i \frac{1}{\Delta V^i}\left(\frac{\partial L}{\partial\phi_i}\delta\phi_i + \frac{\partial L}{\partial\dot{\phi}_i}\delta\dot{\phi}_i\right)\Delta V^i$$

Now, consider the Lagrangian variation using the definition of functional derivative:

$$\delta L = \int d^3\mathbf{x}\left(\frac{\delta L}{\delta\phi}\delta\phi + \frac{\delta L}{\delta\dot{\phi}}\delta\dot{\phi}\right)$$

From the two above equations, we can make the following identifications:

$$\frac{\delta L}{\delta\phi} = \lim_{\Delta V^i\to 0}\frac{1}{\Delta V^i}\frac{\partial L}{\partial\phi_i}$$
$$\frac{\delta L}{\delta\dot{\phi}} = \lim_{\Delta V^i\to 0}\frac{1}{\Delta V^i}\frac{\partial L}{\partial\dot{\phi}_i}$$

In the LHS, while taking functional derivative, $\mathbf{x}$ belongs to the i^{th} cell. Thus, we see that the functional derivative is proportional to the partial derivative. Now, given a Lagrangian, we can write the variation under $\phi \to \phi + \delta\phi$ using the Lagrangian density:

$$\delta L = \int d^3\mathbf{x}\ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_i\phi)}\delta(\partial_i\phi) + \frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\dot{\phi}$$

Integration by parts and ignoring boundaries on the second term, we obtain:

$$\delta L = \int d^3\mathbf{x}\ \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_i\left(\frac{\partial\mathcal{L}}{\partial(\partial_i\phi)}\right)\right)\delta\phi + \frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\dot{\phi}$$

Comparing this with the definition of functional derivative, we obtain:

$$\frac{\delta L}{\delta\phi} = \frac{\partial\mathcal{L}}{\partial\phi} - \partial_i\left(\frac{\partial\mathcal{L}}{\partial(\partial_i\phi)}\right) \qquad \frac{\delta L}{\delta\dot{\phi}} = \frac{\partial\mathcal{L}}{\partial\dot{\phi}}$$

6.2. Hamilton's Equations of Motion

To obtain HEOM, we start with the variation of the total Hamiltonian,

$$\delta H = \int d^3\mathbf{x}\ \delta\big(\pi^\alpha\dot{\phi}_\alpha - \mathcal{L}\big) = \int d^3\mathbf{x}\ \big(\pi^\alpha\delta\dot{\phi}_\alpha + \dot{\phi}_\alpha\delta\pi^\alpha\big) - \delta L$$

Now, from ELEOM:

$$\delta L = \int \frac{\delta L}{\delta\phi_\alpha}\delta\phi_\alpha + \frac{\delta L}{\delta\dot{\phi}_\alpha}\delta\dot{\phi}_\alpha = \int \partial_t\Big(\frac{\delta L}{\delta\dot{\phi}_\alpha}\Big)\delta\phi_\alpha + \pi^\alpha\delta\dot{\phi}_\alpha = \int \dot{\pi}^\alpha\delta\phi_\alpha + \pi^\alpha\delta\dot{\phi}_\alpha$$

Using this, the variation in Hamiltonian becomes:

$$\delta H = \int d^3\mathbf{x}\ (\dot{\phi}_\alpha\delta\pi^\alpha - \dot{\pi}^\alpha\delta\phi_\alpha)$$

Thus, from above, we can identity:

$$\frac{\delta H}{\delta\phi_\alpha} = -\dot{\pi}^\alpha \qquad\qquad \frac{\delta H}{\delta\pi^\alpha} = \dot{\phi}_\alpha$$

Also, previously we saw the relation between functional derivative of the Lagrangian and partial derivative of the Lagrangian density. That could well be extended to the Hamiltonian and we obtain:

$$\dot{\pi}^\alpha = -\frac{\delta H}{\delta\phi_\alpha} = -\frac{\partial\mathcal{H}}{\partial\phi_\alpha} + \partial_i\Big(\frac{\partial\mathcal{H}}{\partial(\partial_i\phi_\alpha)}\Big) \qquad\qquad \dot{\phi}_\alpha = \frac{\delta H}{\delta\pi^\alpha} = \frac{\partial\mathcal{H}}{\partial\pi^\alpha} - \partial_i\Big(\frac{\partial\mathcal{H}}{\partial(\partial_i\pi^\alpha)}\Big)$$

In general, the Hamiltonian density does not depend on $\partial_i\pi^\alpha$, so we have $\dot{\phi}_\alpha = \frac{\partial\mathcal{H}}{\partial\pi^\alpha}$

6.3. Poisson Bracket

In the Hamiltonian formulation of classical mechanics, we defined the Poisson bracket of two quantities A, B by:

$$\{A, B\} = \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}$$

Similarly, for fields also, given two functionals $\mathcal{F}[\phi,\pi]$ and $\mathcal{G}[\phi,\pi]$, we define the Poisson bracket by:

$$\{\mathcal{F},\mathcal{G}\} := \int d^3\mathbf{x}\ \frac{\partial\mathcal{F}}{\partial\phi_\alpha}\frac{\partial\mathcal{G}}{\partial\pi^\alpha} - \frac{\partial\mathcal{F}}{\partial\pi^\alpha}\frac{\partial\mathcal{G}}{\partial\phi_\alpha}$$

Given a function $\mathcal{F}[\phi^\alpha,\pi^\alpha,t]$, note:

$$\begin{aligned}
\dot{\mathcal{F}} &= \partial_t\mathcal{F} + \int d^3\mathbf{x}\ \Big(\frac{\delta\mathcal{F}}{\delta\phi_\alpha}\frac{\partial\phi_\alpha}{\partial t} + \frac{\delta\mathcal{F}}{\delta\pi_\alpha}\frac{\partial\pi_\alpha}{\partial t}\Big) \\
&= \partial_t\mathcal{F} + \int d^3\mathbf{x}\ \Big(\frac{\delta\mathcal{F}}{\delta\phi_\alpha}\frac{\delta H}{\delta\pi^\alpha} - \frac{\delta\mathcal{F}}{\delta\pi_\alpha}\frac{\delta H}{\delta\phi_\alpha}\Big) \qquad \text{(from HEOM)} \\
&= \partial_t\mathcal{F} + \{\mathcal{F}, H\}
\end{aligned}$$

This specifies the *time-evolution* of the functionals using the Poisson bracket. Now, let us consider the Poisson bracket of the fields ϕ_α and π^α. However, since the Poisson bracket was defined for functionals only, there is some problem. This can be resolved by noting that any function can be written as a functional depending on itself as:

$$\phi_\alpha(\mathbf{x},t) = \int d^3\ \mathbf{x}'\phi_\alpha(\mathbf{x}',t)\delta^3(\mathbf{x}-\mathbf{x}')$$

From the definition of the functional derivative, it is easy to note that:

$$\frac{\delta\phi_\alpha(\mathbf{x},t)}{\delta\phi_\beta(\mathbf{x}',t)} = \delta_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{x}')$$

Also, ϕ and π are *independent functionals* and hence,

$$\frac{\delta\pi(\mathbf{x},t)}{\delta\phi(\mathbf{x}',t)} = 0 \qquad \frac{\delta\phi(\mathbf{x},t)}{\delta\pi(\mathbf{x}',t)} = 0 \; \forall \; \mathbf{x},\mathbf{x}',t$$

Now, we can compute the Poisson bracket of the fields themselves and we obtain:

$$\begin{aligned}\{\phi_\alpha(\mathbf{x},t),\pi^\beta(\mathbf{x}',t)\} &= \int d^3\left(\mathbf{x}''\frac{\delta\phi_\alpha(\mathbf{x},t)}{\delta\phi_\mu(\mathbf{x}'',t)}\,\frac{\delta\pi^\beta(\mathbf{x}',t)}{\delta\pi^\mu(\mathbf{x}'',t)} - 0\right) \\ &= \int d^3\left(\mathbf{x}''\delta_{\alpha\mu}\delta_{\beta\mu}\delta^3(\mathbf{x}-\mathbf{x}'')\delta^3(\mathbf{x}'-\mathbf{x}'')\right) \\ &= \delta_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{x}')\end{aligned}$$

We also have the other two Poisson brackets which are trivial:

$$\{\phi_\alpha,\phi_\beta\} = 0 \qquad \{\pi^\alpha,\pi^\beta\} = 0$$

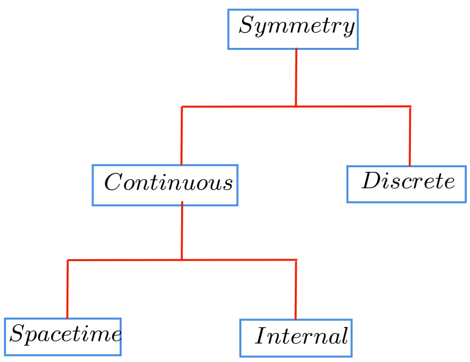
Lecture 07 & 08: Representations of Rotation and Lorentz group

A *group* is a tuple $(G,*)$ which follows certain rules like associativity, existence of identity, inverses, etc. A *Lie Group* [1] is some continuous group with properties so special that it becomes very important (and cool) to study them.

I think a short, vague introduction cannot do justice to these elegant topics which are subjects on their own right and require proper analysis. Hence, I might skip them here and instead try to include them in the link [2] below as and when I read about them!

Lecture 09: Noether's Theorem

To understand Noether's Theorem, let us first understand what symmetry actually is. Symmetry of a system can be thought of as some transformation under which the properties of the system remain invariant.



Discrete symmetries describes non-continuous changes in the system, like *parity* (kind of spatial reflection) and *time-reversal*. *Internal* symmetries include cases where the fields themselves vary while *space-time* symmetry describes the system when the space-time points vary. *Local* symmetries are transformations allowing us to change the system differently at different spacetime points while *global* symmetries act in the same way at every point.

[1] A Lie Group is a group that is also a manifold.
[2] A very basic and nasty introduction can be found here!

Theorem 2 (Noether's Theorem):
For every continuous global symmetry there exists a conserved current J^μ, that is,

$$\partial_\mu J^\mu = 0$$

Proof. Note that, under any symmetry transformation, we required our equations of motion to be invariant and hence require our action $\mathcal{S}$ to be 'invariant' since action is a fundamental quantity to our system from which we derived our EOM. Note that while deriving the equations of motions, we ignored the boundary term of the action. Thus, the action can vary by some arbitrary boundary term. Using this, we define a *Noether symmetry* to be a transformation which keeps the action invariant upto an integral of a total divergence, that is,

$$\delta\mathcal{S} = \int\limits_{\partial\mathcal{R}} d^4\mathbf{x}\ \partial_\mu K^\mu \implies \delta\mathcal{L} \equiv \epsilon\ \partial_\mu K^\mu$$

The implication comes from the fact that divergence should be in first order with respect to some infinitesimal variation ϵ. Let us consider the case of *internal symmetries* first.

Consider a variation in the field $\phi = \phi_0 + \delta\phi$. Then,

$$\delta\mathcal{L} = \left.\frac{\partial\mathcal{L}}{\partial\phi}\right|_{\phi_0}\delta\phi + \left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\phi_0}\delta(\partial_\mu\phi) = \left.\frac{\partial\mathcal{L}}{\partial\phi}\right|_{\phi_0}\delta\phi + \left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\phi_0}\partial_\mu(\delta\phi)$$

Now, suppose there exists a solution $\widetilde{\phi}$ to the equation of motion of the system. Since the above equation is valid for any ϕ_0, it is also valid for $\widetilde{\phi}$ and hence, we get:

$$\begin{aligned}
\delta\mathcal{L} &= \left.\frac{\partial\mathcal{L}}{\partial\phi}\right|_{\widetilde{\phi}}\delta\phi + \left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\partial_\mu(\delta\phi) \\
&= \partial_\mu\left.\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right|_{\widetilde{\phi}}\delta\phi + \left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\partial_\mu(\delta\phi) \quad \text{(using ELEOM)} \\
&= \partial_\mu\left(\left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\delta\phi\right) \quad \text{(using product rule)}
\end{aligned}$$

Let the variation of the Lagrangian be a global symmetry, denoted by, $\delta_\alpha\phi$. Under this symmetry, we want our action $\mathcal{S}$ to change atmost by a boundary term and hence we have,

$$\partial_\mu\left(\left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\delta\phi\right) = \partial_\mu K^\mu \quad\implies\quad \partial_\mu\left(\left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\delta\phi - K^\mu\right) = 0$$

Let us define the quantity in the bracket by J^μ which is called the *Noether current*. Thus,

$$J^\mu := \left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right|_{\widetilde{\phi}}\delta\phi - K^\mu \tag{1}$$

from where we get $\partial_\mu J^\mu = 0 \implies$ Noether current is conserved.

The above proof worked for internal symmetries where the field itself was varied. Now, let us look into the *spacetime* symmetry, where the spacetime points are varied. For simplicity, we will deal with scalar fields only. Let us consider an infinitesimal coordinate transformation,

$$x'^\mu = x^\mu + \delta x^\mu$$

The variation in the field can be given by:

$$\delta\phi = \phi'(\mathbf{x}') - \phi(\mathbf{x}) = \phi'(x'^{\mu}) - \phi'(x^{\mu}) + \phi'(x^{\mu}) - \phi(x^{\mu}) = \delta_0\phi(x^{\mu}) + \phi'(x'^{\mu}) - \phi'(x^{\mu})$$

where we have defined $\delta_0\phi(x^{\mu}) := \phi'(x^{\mu}) - \phi(x^{\mu})$. We did this to distinguish between variation of the dependent and independent quantitites. Using Taylor expansion, we have:

$$\phi'(x'^{\mu}) = \phi'(x^{\mu} + \delta x^{\mu}) = \phi'(x^{\mu}) + \frac{\partial\phi'}{\partial x^{\mu}}\delta x^{\mu} \implies \phi'(x'^{\mu}) - \phi'(x^{\mu}) = \partial_{\mu}\phi'\delta x^{\mu}$$

Now, $\partial_{\mu}\phi'\delta x^{\mu} \approx \partial_{\mu}\phi\delta x^{\mu}$, since if we do the Taylor expansion of $\delta_0\phi(x'^{\mu})$, we would obtain a term $\partial_{\nu}\delta_0$ which is a variation of second order and hence, can be neglected. Thus, we obtain:

$$\phi'(x'^{\mu}) - \phi'(x^{\mu}) = \partial_{\mu}\phi\delta x^{\mu} \implies \delta\phi = \delta_0\phi(x^{\mu}) + \partial_{\mu}\phi\delta x^{\mu}$$

The action is not in general invariant under both changes in coordinates and fields. Indeed, the volume element (integral measure) $\mathrm{d}^4\mathbf{x}$ transforms as:

$$\mathrm{d}^4\mathbf{x}' = J\ \mathrm{d}^4\mathbf{x}$$

where J is the determinant of the Jacobian matrix. Now, we find some expression for the Jacobian:

$$x'^{\mu} = x^{\mu} + \delta x^{\mu} \implies \partial_{\nu}x'^{\mu} = \delta^{\mu}{}_{\nu} + \partial_{\nu}(\delta x^{\mu})$$

Now, we use an identity $\det(\mathbb{1} + \epsilon A) = 1 + \epsilon \operatorname{Tr} A + \mathcal{O}(\epsilon^2)$ on the above equation to get an approximation for the Jacobian as:

$$J = \det\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right) = \det(\delta^{\mu}{}_{\nu} + \partial_{\nu}(\delta x^{\mu})) \approx (1 + \partial_{\nu}(\delta x^{\mu})) \implies \delta J = \partial_{\mu}(\delta x^{\mu})$$

Now, we consider the variation of the action itself and we see:

$$\begin{aligned}\delta\mathcal{S} &= \int \delta(\mathrm{d}^4\mathbf{x})\ \mathcal{L} + \int \mathrm{d}^4\mathbf{x}\ \delta\mathcal{L} \\ &= \int \mathrm{d}^4\mathbf{x}\ \partial_{\mu}\delta x^{\mu}\mathcal{L} + \int \mathrm{d}^4\mathbf{x}\ \delta\mathcal{L} \\ &= \int \mathrm{d}^4\mathbf{x}\ \left(\partial_{\mu}\delta x^{\mu}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\phi}\delta_0\phi + \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\delta_0(\partial_{\mu}\phi) + \frac{\partial\mathcal{L}}{\partial x^{\mu}}\delta x^{\mu}\right)\end{aligned}$$

We had written the variation of the Lagrangian with respect to both dependent and independent variables. Now the first and last term can be written together using product rule and hence we have:

$$\begin{aligned}\delta\mathcal{S} &= \int \mathrm{d}^4\mathbf{x}\ \left(\partial_{\mu}(\delta x^{\mu}\mathcal{L}) + \frac{\partial\mathcal{L}}{\partial\phi}\delta_0\phi + \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\partial_{\mu}\delta_0\phi\right) \\ &= \int \mathrm{d}^4\mathbf{x}\ \left(\partial_{\mu}(\delta x^{\mu}\mathcal{L}) + \frac{\partial\mathcal{L}}{\partial\phi}\delta_0\phi + \partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\delta_0\phi\right) - \partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\right)\delta_0\phi\right) \\ &= \int \mathrm{d}^4\mathbf{x}\ \partial_{\mu}\left(\delta x^{\mu}\mathcal{L} + \left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\delta_0\phi\right)\right) + \delta_0\phi\left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\right)\right)\end{aligned}$$

For our field satisfying the equation of motion, the last term drops and we have our variation of the action and this should be equal to the boundary term, hence we have:

$$\partial_{\mu}\left(\delta x^{\mu}\mathcal{L} + \left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\delta_0\phi\right)\right) = \partial_{\mu}K^{\mu}$$

In the case for multiple fields, the *Noether current* turns out to be:

$$J^{\mu} := \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_i)}\delta_0\phi_i + \delta x^{\mu}\mathcal{L} - K^{\mu} \qquad \partial_{\mu}J^{\mu} = 0 \quad \square$$

The above condition is called *current conservation* which we had also seen earlier for internal symmetries. However, the second term, which is purely an artifact of the spacetime symmetry, did not occur there in the conserved current for internal symmetries. Associated with each current, we can define the charge enclosed in a volume V as:

$$Q_V := \int_V \mathrm{d}^3\mathbf{x}\ J^0(\mathbf{x},t) = \int_V \mathrm{d}^3\mathbf{x}\ \rho$$

Then, we can write the conservation equation as:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \boldsymbol{\nabla}\cdot\mathbf{J} = 0 \implies \frac{\mathrm{d}Q}{\mathrm{d}t} = -\int_V \mathrm{d}^3\mathbf{x}\ \boldsymbol{\nabla}\cdot\mathbf{J} = -\int_{\partial V} \mathbf{J}\cdot\mathrm{d}\mathbf{S}$$

The last equality comes from Gauss' law and ∂V is the boundary of the region. If we extend the volume V to cover the entire space and assume that at spatial infinities, current is zero, then we have $\frac{\mathrm{d}Q}{\mathrm{d}t} = 0$ and hence, current is conserved. In most cases, we would have K^μ to be zero (and thus, let us take it to be zero). Let us see some examples for Noether's theorem:

9.1. Examples: Spacetime Symmetries

Let us start with some examples of Noether's theorem where spactime symmetry is explicitly there. We will consider translations and rotations.

9.1.1. Translational Invariance

Classically, we know that translational invariance leads to *energy conservation*. Let us see what happens in case of fields. We thus consider an infinitesimal global translation:

$$x^\mu \to x'^\mu = x^\mu + \epsilon a^\mu$$

Then $\delta x^\mu = \epsilon a^\mu$. As calculated before, the variation in the Lagrangian is given by:

$$\delta\mathcal{L} = \partial_\mu\left(\delta x^\mu \mathcal{L} + \delta_0\phi_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)$$

Note that $\delta x^\mu = \epsilon a^\mu$ and $\delta_0\phi_i = \phi_i'(x^\mu) - \phi_i(x^\mu)$. Note that the new field at x^μ is just the previous field at $x^\mu - \epsilon a^\mu$ and hence we have from a first-order expansion that $\delta_0\phi_i = -\epsilon a^\nu\partial_\nu\phi$. Substituting this above, we obtain:

$$\begin{aligned}
\partial_\mu\left(\delta x^\mu \mathcal{L} + \delta_0\phi_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) &= \partial_\mu\left(\epsilon a^\mu \mathcal{L} - \epsilon a^\nu\partial_\nu\phi_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= \partial_\mu\left(\epsilon a^\nu \delta^\mu{}_\nu \mathcal{L} - \epsilon a^\nu\partial_\nu\phi_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= -\epsilon a^\nu \partial_\mu\left(\partial_\nu\phi_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)} - \delta^\mu{}_\nu\mathcal{L}\right)
\end{aligned}$$

From the above, we can thus define the conserved current as:

$$T^\mu{}_\nu := \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial_\nu\phi_i - \delta^\mu{}_\nu\mathcal{L}$$

Now, note the following:

$$T^{\mu\nu} = g^{\alpha\nu}T^\mu{}_\alpha = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}g^{\alpha\nu}\partial_\alpha\phi_i - g^{\alpha\nu}\delta^\mu{}_\alpha\mathcal{L} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial^\nu\phi_i - g^{\mu\nu}\mathcal{L}$$

The above quantity can be shown to be a tensor and is called the *energy-momentum stress tensor*. Why?

Note that for each value of ν we have a current four vector. Let us calculate the charge for the time-translation invariance part, that is, $\nu = 0$. From this, we have:

$$P^0 = \int \mathrm{d}^3\mathbf{x}\ T^{00} = \int \mathrm{d}^3\mathbf{x}\ \left(\frac{\partial \mathcal{L}}{\partial(\partial_0\phi_i)}\partial^0\phi_i - g^{00}\mathcal{L}\right) = \int \mathrm{d}^3\mathbf{x}\ \left(\pi^i\dot{\phi}_i - \mathcal{L}\right) = \int \mathrm{d}^3\mathbf{x}\ \mathcal{H} = H$$

We see that corresponding to the conserved current from time translation invariance, we have a conserved charge $P^0 = H$. Thus, the total Hamiltonian is conserved under time translation and hence our classical intuition worked here as well!

Now, let us see what happens for space translations, for which we consider $\nu \equiv j = 1, 2, 3$. Then we have:

$$P^j = \int \mathrm{d}^3\mathbf{x}\ T^{0j} = \int \mathrm{d}^3\mathbf{x}\ \left(\frac{\partial \mathcal{L}}{\partial(\partial_0\phi_i)}\partial^j\phi_i - \cancelto{0}{g^{0j}}\ \mathcal{L}\right) = \int \mathrm{d}^3\mathbf{x}\ \left(\pi^i\partial^j\phi_i\right)$$

The last thing is the j^{th} component of the total momentum which is conserved in this case. Thus, we see that for spatial translation invariance, the total momentum is conserved, which once again matches with the classical analog.

9.1.2. Lorentz Invariance

An infinitesimal Lorentz transformation (which covers both *boosts* and *rotations*) can be written as:

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

where $\omega^\mu{}_\nu$ is an infinitesimal deviation from the identity. Now, we know for any Lorentz transformation,

$$\Lambda^\mu{}_\tau \eta^{\tau\sigma} \Lambda^\nu{}_\sigma = \eta^{\mu\nu} \iff \Lambda^{\mathrm{T}}\eta\Lambda = \eta$$

If we put the infinitesimal transformation in the above equation and retain only terms upto first order in ω, we would have $\omega^{\mu\nu} = \omega^{\nu\mu}$, that is, the infinitesimal deviation is an anti-symmetric matrix. The spacetime coordinates transforms as:

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu = \delta^\mu{}_\nu x^\nu + \omega^\mu{}_\nu x^\nu = x^\mu + \omega^\mu{}_\nu x^\nu \implies \delta x^\mu = \omega^\mu{}_\nu x^\nu$$

And the variation in ϕ is given as:

$$\delta_0\phi_i = \phi'(x^\mu) - \phi(x^\mu) = \phi(x^\mu - \omega^\mu{}_\nu x^\nu) - \phi(x^\mu) = -\omega^\mu{}_\nu x^\nu \partial_\mu \phi_i$$

Now, we are ready to put in the variation of the Lagrangian:

$$\begin{aligned}
\partial_\mu\left(\delta x^\mu \mathcal{L} + \delta_0\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) &= \partial_\mu\left(\omega^\mu{}_\nu x^\nu \mathcal{L} - \omega^\sigma{}_\nu x^\nu \partial_\sigma\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= \partial_\mu\left(\eta^{\mu\rho}\omega_{\rho\nu} x^\nu \mathcal{L} - \eta^{\sigma\lambda}\omega_{\lambda\nu} x^\nu \partial_\sigma\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= \partial_\mu\left(\eta^{\mu\rho}\omega_{\rho\nu} x^\nu \mathcal{L} - \omega_{\lambda\nu} x^\nu \partial^\lambda\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= \partial_\mu\left(\eta^{\mu\rho}\omega_{\rho\nu} x^\nu \mathcal{L} - \omega_{\rho\nu} x^\nu \partial^\rho\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) \\
&= \partial_\mu\left(-\omega_{\rho\nu} x^\nu \times \left(-\eta^{\mu\rho}\mathcal{L} + \partial^\rho\phi_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)\right) \\
&= \partial_\mu(-\omega_{\rho\nu} x^\nu T^{\rho\mu}) \\
&= \partial_\mu(\omega_{\nu\rho} x^\nu T^{\rho\mu})
\end{aligned}$$

Since the current must be conserved for all the six elements of the anti-symmetric matrix (since anti-symmetric so there are total $\frac{4\times(4-1)}{2} = 6$ independent entries), the part of $x^\nu T^{\rho\mu}$ which is anti-symmetric in ρ and μ must be conserved. After antisymmetrising and considering current for each component, we have:

$$M^{\nu\mu\rho} := x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} \qquad \partial_\mu M^{\nu\mu\rho} = 0$$

Let us now see the conserved charges for this case. Corresponding to spatial rotations, we have:

$$\mathcal{J}^{ij} = \int \mathrm{d}^3\mathbf{x}\ M^{i0j} = \int \mathrm{d}^3\mathbf{x}\ (x^i T^{0j} - x^j T^{0i})$$

This can be related to the angular-momentum since T^{0i} was related to P^i, the i^{th} component of the momentum. Hence, rotational invariance leads to the *conservation of angular momentum.* Let us see what happens when we give *boosts.*

$$\mathcal{J}^{0j} = \int \mathrm{d}^3\mathbf{x}\ M^{00j} = \int \mathrm{d}^3\mathbf{x}\ (x^0 T^{0j} - x^j T^{00})$$

As $t = x^0$ and we know that the current $\mathcal{J}^{0j}$ must be conserved, we have:

$$\frac{\mathrm{d}\mathcal{J}^{0j}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\mathbf{x}\ (x^0 T^{0j} - x^j T^{00}) = \int \mathrm{d}^3\mathbf{x}\ T^{0j} + \int \mathrm{d}^3\mathbf{x}\ t\ \partial_t T^{0j} - \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\mathbf{x}\ x^j T^{00} = 0$$

In terms of P^j, the above equation can be written as:

$$P^j + t\frac{\mathrm{d}P^j}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\mathbf{x}\ x^j T^{00} = 0$$

Previously we had seen that P^j was the conserved charge corresponding to spatial translation and hence P^j is a constant, which gives us:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\mathbf{x}\ x^j T^{00} = \mathcal{C}\ (const.)$$

The above relation can be seen as the centre of energy (or mass) of the field travels with constant velocity as T^{00} can be treated as the Hamiltonian ($\approx$ energy) density.

9.2. Examples: Internal Symmetries

Let us now look at internal symmetries, where the spacetime coordinates themselves do not change but the fields change due to some global transformations.

9.2.1. Global U(1) Symmetry of Complex Scalar Field

Consider a complex scalar field ϕ with the following Lagrangian:

$$\mathcal{L} = (\partial_\mu\phi)(\partial^\mu\phi^*) - m^2\phi\phi^*$$

Let us consider a $U(1)$ transformation [1]. Note that elements of the group $U(1)$ are represented by $e^{i\alpha}$, that is, they lie on the unit circle. Then, for an infinitesimal change $\psi \to e^{i\alpha}\psi$, $e^{i\alpha} = 1 + i\alpha + \mathcal{O}(\alpha^2)$[2] and we have:

$$\phi' = (1 + i\alpha)\phi \implies \delta_\alpha\phi = i\alpha$$
$$\phi^{*\prime} = (1 - i\alpha)\phi^* \implies \delta_\alpha\phi^* = -i\alpha$$

Under this transformation, we see that the Lagrangian is invariant, that is,

$$\mathcal{L}' = (\partial_\mu\phi')(\partial^\mu\phi'^*) - m^2\phi'\phi'^* = e^{i\alpha}e^{-i\alpha}(\partial_\mu\phi)(\partial^\mu\phi'^*) - e^{i\alpha}e^{-i\alpha}m^2\phi'\phi'^* = \mathcal{L}$$

From this, we can take $K^\mu = 0$. Using the formula for the conserved current, we have:

$$J^\mu = i\alpha\phi(\partial^\mu\phi^*) - i\alpha\phi^*(\partial^\mu\phi) \sim i(\phi\partial^\mu\phi^* - \phi^*\partial^\mu\phi) \tag{2}$$

The conserved charge is then given by the volume integral:

$$Q = \int \mathrm{d}^3\mathbf{x}\ J^0 = -i\int \mathrm{d}^3\mathbf{x}\ (\phi\dot{\phi}^* - \dot{\phi}\phi^*)$$

[1] This is an *internal* symmetry as spacetime points are not varied.
[2] This is a *global* symmetry as α does not depend on the spacetime coordinates

Lecture 10 & 11: Field Quantisation

So far we had worked with classical fields. In the introduction, we saw that, quantum theories due to decoherence, transforms to classical theories, with an inevitable loss of information. Thus, finding the quantum theory leading to the classical model (working backwards) is not *unique*. More than one quantum field theory might have the same classical limit. Thus, while doing the quantisation process, we always have to make an *educated guess*!

Now, let us make the KG field *quantum* by quantising the canonical coordinates, similar to as we did for the harmonic oscillator.

We have our classical field $\phi(\mathbf{x}, t)$ and its canonical conjugate momentum $\pi(\mathbf{x}, t)$ which become operators $\hat{\phi}(x)$ and $\hat{\pi}(x)$[1]. The classical Poisson bracket changes to the commutator, that is, $\{\cdot, \cdot\} \to i[\cdot, \cdot]$.

A quantum treatment of a classical Hamiltonian is sometimes messy since we have to maintain the Hermiticity condition and there can be multiple ways to do that. However, almost always, we just put the hats on appropriate places and wish we are correct!

Classically, we had $\{\phi(\mathbf{x}), \pi(\mathbf{y})\} = \delta^{(3)}(\mathbf{x} - \mathbf{y})$ and by the same logic:

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

Let us take their Fourier transforms,

$$\widetilde{\phi}(\mathbf{k}, t) = \int \mathrm{d}^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \phi(\mathbf{x}, t)$$
$$\widetilde{\pi}(\mathbf{k}, t) = \int \mathrm{d}^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \pi(\mathbf{x}, t)$$

Let us now consider the Klein-Gordon equation,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2 \iff (\partial_\mu\partial^\mu + m^2)\phi = 0$$

Expanding the field in Fourier space, we would obtain the following:

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

Note that this is the equation for a classical harmonic oscillator with frequency $\omega_\mathbf{p} = \sqrt{|\mathbf{p}|^2 + m^2}$ and the general solution $\widetilde{\phi}(\mathbf{x}, t)$ will be a linear combination of $e^{i\omega_\mathbf{p}}$ and $e^{-i\omega_\mathbf{p}}$. Thus we have,

$$\widetilde{\phi}(\mathbf{p}, t) = A(\mathbf{p})e^{i\omega_\mathbf{p}t} + B(\mathbf{p})e^{-i\omega_\mathbf{p}t} \qquad \omega_\mathbf{p} = \sqrt{|\mathbf{p}|^2 + m^2}$$

Now, in the quantum domain, $A(\mathbf{p})$ and $B(\mathbf{p})$ will become the creation and annihilation operators $a_\mathbf{p}$ and $a^\dagger_\mathbf{p}$ and in that spirit, let us define the field operators in terms of the annihilation and creation operators for each Fourier mode. Note that, for simplicity we define the field operators at $t = 0$.

$$\hat{\phi}(\mathbf{x}) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left(\hat{a}_\mathbf{p}e^{i\mathbf{p}\cdot\mathbf{x}} + \hat{a}^\dagger_\mathbf{p}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)$$
$$\hat{\pi}(\mathbf{x}) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_\mathbf{p}}{2}}\left(\hat{a}_\mathbf{p}e^{i\mathbf{p}\cdot\mathbf{x}} - \hat{a}^\dagger_\mathbf{p}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)$$

where $\omega_\mathbf{p} = \sqrt{|\mathbf{p}|^2 + m^2}$ is the frequency of the mode with momentum $\mathbf{p}$. Replacing $\mathbf{p}$ with $-\mathbf{p}$ in the second term, the above can also be written in a more compact way as:

$$\hat{\phi}(\mathbf{x}) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left(\hat{a}_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}$$
$$\hat{\pi}(\mathbf{x}) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_\mathbf{p}}{2}}\left(\hat{a}_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}$$

[1]Note that we will write four vectors with non-bold letter, as x will denote the four components of the four-vector. To denote the three vectors, we will write in bold-face, that is, $\mathbf{x}$

Let us check the commutation relation:

$$\left[\hat{\phi}(\mathbf{x}), \hat{\pi}(\mathbf{y})\right] = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{(-i)}{2}\sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{p}'}}}\left(\left[\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}, \hat{a}_{\mathbf{p}'} - \hat{a}^\dagger_{-\mathbf{p}'}\right]\right)e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{p}'\cdot\mathbf{y})} = i\delta^3{}_{\mathbf{x}-\mathbf{y}}$$

Note that the above relation can be satisfied if we take:

$$\left[\hat{a}_{\mathbf{p}}, \hat{a}^\dagger_{\mathbf{p}'}\right] = (2\pi)^3\delta^3{}_{\mathbf{p}-\mathbf{p}'}\mathbb{1}$$

Now, let us do a horrenduous calculation, which will then lead to a horrenduous result and after some reasoning based on physical grounds, we will obtain a cute result. From the Hamiltonian density, we have the total Hamiltonian as:

$$\hat{H} = \int \mathrm{d}^3\mathbf{x}\ \mathcal{H} = \int \mathrm{d}^3\mathbf{x}\left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\nabla\hat{\phi})^2 + \frac{1}{2}m^2\hat{\phi}^2\right)$$

Let us calculate this term by term. The first term gives us:

$$\begin{aligned}
\int \mathrm{d}^3\mathbf{x}\ \pi^2(\mathbf{x}) &= \int \mathrm{d}^3\mathbf{x}\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}{-2}(\hat{a}_{\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} - \hat{a}^\dagger_{-\mathbf{p}'})e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\mathrm{d}^3\mathbf{p}'\frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}{-2}(\hat{a}_{\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} - \hat{a}^\dagger_{-\mathbf{p}'})\delta^3{}_{\mathbf{p}+\mathbf{p}'} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left[-\frac{\omega_{\mathbf{p}}}{2}\right](\hat{a}_{\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{-\mathbf{p}} - \hat{a}^\dagger_{\mathbf{p}}) \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left[-\frac{\omega_{\mathbf{p}}}{2}\right](\hat{a}_{\mathbf{p}}\hat{a}_{-\mathbf{p}} - \hat{a}_{\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}})
\end{aligned}$$

From the last term, we have:

$$\begin{aligned}
\int \mathrm{d}^3\mathbf{x}\ \phi^2(\mathbf{x}) &= \int \mathrm{d}^3\mathbf{x}\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} + \hat{a}^\dagger_{-\mathbf{p}'})e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\mathrm{d}^3\mathbf{p}'\frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} + \hat{a}^\dagger_{-\mathbf{p}'})\delta^3{}_{\mathbf{p}+\mathbf{p}'} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left[\frac{1}{2\omega_{\mathbf{p}}}\right](\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{\mathbf{p}}) \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left[\frac{1}{2\omega_{\mathbf{p}}}\right](\hat{a}_{\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}_{\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}})
\end{aligned}$$

Now comes the second term. For this, note that we have $\nabla(e^{i\mathbf{k}\cdot\mathbf{x}}) = i\mathbf{k}(e^{i\mathbf{k}\cdot\mathbf{x}})$. Using this, we obtain:

$$\begin{aligned}
\int \mathrm{d}^3\mathbf{x}\ (\nabla\phi)\cdot(\nabla\phi) &= \int \mathrm{d}^3\mathbf{x}\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}\left[i\mathbf{p}(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})e^{i\mathbf{p}\cdot\mathbf{x}}\right]\cdot\left[i\mathbf{p}'(\hat{a}_{\mathbf{p}'} + \hat{a}^\dagger_{-\mathbf{p}'})e^{i\mathbf{p}'\cdot\mathbf{x}}\right] \\
&= -\int \mathrm{d}^3\mathbf{x}\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}\left[(\mathbf{p}\cdot\mathbf{p}')(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} + \hat{a}^\dagger_{-\mathbf{p}'})\right]e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \\
&= -\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\mathrm{d}^3\mathbf{p}'\frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}\left[(\mathbf{p}\cdot\mathbf{p}')(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{\mathbf{p}'} + \hat{a}^\dagger_{-\mathbf{p}'})\right]\delta^3{}_{\mathbf{p}+\mathbf{p}'} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}\left[(\mathbf{p}^2)(\hat{a}_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}})(\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{\mathbf{p}})\right] \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathbf{p}^2}{2\omega_{\mathbf{p}}}\left(\hat{a}_{\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}_{\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}}\right)
\end{aligned}$$

Adding all the terms with correct pre-factors, we obtain the following:

$$\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\Big\{\Big[-\frac{\omega_\mathbf{p}}{4}\Big]\left(\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}} - \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}\right)\Big\} + \Big\{\Big[\frac{m^2}{4\omega_\mathbf{p}}\Big]\left(\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}} + \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}\right)\Big\}$$

$$+ \Big\{\frac{\omega_\mathbf{p}^2 - m^2}{4\omega_\mathbf{p}}\left(\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}} + \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}\right)\Big\}$$

$$= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\omega_\mathbf{p}}{4}\Big(\cancel{\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}}} + \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} + \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}} - \cancel{\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}}} + \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} - \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}}\Big)$$

$$+ \frac{m^2}{4\omega_\mathbf{p}}\Big(\cancel{\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}}} + \cancel{\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p}} + \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}}} + \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}} - \cancel{\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}}} - \cancel{\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p}} - \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}}} - \cancel{\hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}}\Big)$$

$$= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\omega_\mathbf{p}}{2}\left(\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}}\right) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\omega_\mathbf{p}}{2}\left(\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}\right)$$

Now, using the commutator as obtained before, we get:

$$\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\omega_\mathbf{p}}{2}\left((2\pi)^3{\delta^3}_0 + 2\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}\right) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\omega_\mathbf{p}\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p} + \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\omega_\mathbf{p}}{2}{\delta^3}_0$$

This is the end of the horrenduous calculation and now we discuss the horrenduous term (the second term) which was obtained as a result of applying the commutator. Since this contains a delta-function (distribution?) and the integral is over all momentum, this definitely cannot be a finite quantity, hence it diverges. This term basically represents the zero-point energies of the harmonic oscillators (since there are infinite number of harmonic oscillators, we have a divergent zero-point energy). The fortunate thing is that this is a constant.

When physically measuring energies, only differences can be calculated and hence this term gets cancelled and thus, no divergence actually happens.[1]

Thus, in all practicality, we can rescale our energy and we obtain the Hamiltonian as:

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

Now, let us see the commutator of the ladder operators with the Hamiltonian. We have:

$$\left[\hat{H}, \hat{a}^\dagger_\mathbf{p}\right] = \int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\left[\hat{a}^\dagger_\mathbf{p'}\hat{a}_\mathbf{p'}, \hat{a}^\dagger_\mathbf{p}\right] = \int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\hat{a}^\dagger_\mathbf{p'}\left[\hat{a}_\mathbf{p'}, \hat{a}^\dagger_\mathbf{p}\right] = \int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\hat{a}^\dagger_\mathbf{p'}(2\pi)^3{\delta^3}_{\mathbf{p}-\mathbf{p'}} = \omega_\mathbf{p}\hat{a}^\dagger_\mathbf{p} \quad (3)$$

$$\left[\hat{H}, \hat{a}_\mathbf{p}\right] = \int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\left[\hat{a}^\dagger_\mathbf{p'}\hat{a}_\mathbf{p'}, \hat{a}_\mathbf{p}\right] = \int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\left[\hat{a}^\dagger_\mathbf{p'}, \hat{a}_\mathbf{p}\right]\hat{a}_\mathbf{p'} = -\int \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\omega_\mathbf{p'}\hat{a}_\mathbf{p'}(2\pi)^3{\delta^3}_{\mathbf{p}-\mathbf{p'}} = -\omega_\mathbf{p}\hat{a}_\mathbf{p} \quad (4)$$

11.1. Conserved Charges

Note that we already found the Hamiltonian which is one of the conserved charges corresponding to T^{00}. Let us now focus on the momentum, that is, $\hat{P}^j$ corresponding to T^{0j}

$$\hat{P}^j = \int \mathrm{d}^3\mathbf{x}\ \hat{T}^{0j} = -\int \mathrm{d}^3\mathbf{x}\ \hat{\pi}\partial_j\hat{\phi}$$

Note that, while changing from the classical expression to quantum by 'putting hats', we had several choices for the ordering of $\hat{\pi}$ and $\hat{\phi}$ but here, we specify the one which works!

We now have two options, either to do the calculation again or just specify the result of this calculation

[1] This infinite term is sometimes attributed to the *cosmological constant*. It also has some role in the *Casimir effect*.

since this had already been done by stalwarts before. We choose the former and thus, we have:

$$\begin{aligned}\int \mathrm{d}^3\mathbf{x}\ \hat{\pi}\partial_j\hat{\phi} &= \int \mathrm{d}^3\mathbf{x}\ \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_\mathbf{p}}{2}}\frac{1}{\sqrt{2\omega_{\mathbf{p'}}}}\left(\hat{a}_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\right)ip'_j\left(\hat{a}_{\mathbf{p'}} + \hat{a}^\dagger_{-\mathbf{p'}}\right)e^{i(\mathbf{p}+\mathbf{p'})\cdot\mathbf{x}} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\sqrt{\frac{\omega_\mathbf{p}}{2}}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left(\hat{a}_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\right)p'_j\left(\hat{a}_{-\mathbf{p}} + \hat{a}^\dagger_\mathbf{p}\right) \\ &= \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}p_j(\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}} + \hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p})\end{aligned}$$

Now, note the following:

$$\mathcal{I} = \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}p_j(\hat{a}_\mathbf{p}\hat{a}_{-\mathbf{p}} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}^\dagger_\mathbf{p}) \quad \overset{\mathbf{p}\to-\mathbf{p}}{=} \quad \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(-p_j)(\hat{a}_{-\mathbf{p}}\hat{a}_\mathbf{p} - \hat{a}^\dagger_\mathbf{p}\hat{a}^\dagger_{-\mathbf{p}}) = -\mathcal{I}$$

The last equality comes since the creation and annihilation operators commute for different values of the momentum. Thus, this integral is zero and the expression for the total momentum becomes:

$$\hat{P}^j = \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}p_j\left(\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} - \hat{a}^\dagger_{-\mathbf{p}}\hat{a}_{-\mathbf{p}}\right) = \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}p_j\left(\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p} + \hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}\right) = \frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}p_j\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}$$

In the last term, we introduced the commutator and removed the constant infinity term.

11.2. Normal Ordering

The way we are doing, like removing the infinities, is a common thing in QFT. This purely results from the fact the things commute in the classical case and hence there is always an order ambiguity when we move to quantum.

For example, if in classical scenario, we have some term like xp, when elevating (idk if its really elevating) these things to operators, we can have multiple possibilities, viz. $\hat{x}\hat{p}$, $\hat{p}\hat{x}$ or even $\frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$ (placing things symmetrically). Each one would present some additional constant on the final result, since the commutator of $\hat{p}$ and $\hat{x}$ is just a constant.

To be consistent, we generally use *normal ordering*, that is, keeping all the creation operators to the left. For that we define the *normal ordering operator*, $\widehat{\mathcal{N}}$. For example,

$$\widehat{\mathcal{N}}[\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}] = \hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p} \qquad \widehat{\mathcal{N}}[\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p}] = \hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p} \qquad \widehat{\mathcal{N}}[\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}\hat{a}_\mathbf{p}\hat{a}^\dagger_\mathbf{p}] = \hat{a}^\dagger_\mathbf{p}\hat{a}^\dagger_\mathbf{p}\hat{a}_\mathbf{p}\hat{a}_\mathbf{p} \quad \ldots$$

The $\widehat{\mathcal{N}}$ is not an operator in the sense that is doesnt act on a state to provide some new information. Instead, normal ordering is an interpretation that we use to eliminate the meaningless infinities that occur in field theories. So the rule goes that 'if you want to tell someone about a string of operators in quantum field theory, you have to normal order them first'. Thus, in operators like $\hat{P}$ and $\hat{H}$, in the final expression, we don't use the commutator but the normal operator operator, which will then completely subdue the issue of the appearance of Dirac-delta.

11.3. Constructing the Hilbert Space

The entire basis for this is the harmonic oscillator. In quantum SHO, we started with the vacuum and then acted creation operator successively on the vacuum to create excited states. Here, we will also do the same. Let us define the vacuum state $|\Omega\rangle$ such that,

$$\hat{a}_\mathbf{p}|\Omega\rangle = 0 \ \forall\ \mathbf{p}$$

With such a definition, on physical grounds also, we obtain,

$$\hat{H}|\Omega\rangle = 0 \implies \text{Ground state has zero energy}$$

The vacuum can be written explicitly as:

$$|\Omega\rangle = \bigotimes_{\mathbf{p}} |\Omega\rangle_{\mathbf{p}} \qquad |\Omega\rangle_{\mathbf{p}} \to \text{vac. of SHO with mom. } \mathbf{p}$$

The Hilbert space is then built as the space generated by all finite linear combinations of vectors of the form:

$$|\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle \equiv \hat{a}^\dagger_{\mathbf{p}_1} \hat{a}^\dagger_{\mathbf{p}_2} \dots \hat{a}^\dagger_{\mathbf{p}_n} |\Omega\rangle$$

Till now, this is not exactly the Hilbert space, since we did not account for the norm of the vectors, where a huge subtlely lies.

Note that $|\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle$ is an eigenstate of both $\hat{H}$ and $\hat{P}$ since,

$$\hat{H} |\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle = (\omega_{\mathbf{p}_1} + \dots + \omega_{\mathbf{p}_n}) |\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle$$
$$\hat{P} |\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle = (\mathbf{p}_1 + \dots + \mathbf{p}_n) |\mathbf{p}_1, \mathbf{p}_2, \dots \mathbf{p}_n\rangle$$

Since the creation operators commute, the multi-particle states are symmetric under exchange of particles, that is, the particles follow *Bose-Einstein* statistics.

For a general many-particle system, this 'pre'-Hilbert space is termed as *Fock space* $\mathfrak{F}$ and is formally defined with respect to direct-sums,

$$\mathfrak{F} = \bigoplus_{n=0}^{\infty} \mathscr{H}^{\otimes n}$$

where $\mathscr{H}^{\otimes n}$ is the n-times tensor product of the Hilbert space. In simple terms, $\mathscr{H}^{\otimes 0}$ is the space spanned by the vacuum state, $\mathscr{H}^{\otimes 1}$ is the space spanned by $|\mathbf{p}\rangle$. For $n \geq 2$, we have to consider the fermionic and bosonic properties too, that is, in our case, since these are bosons, $\mathscr{H}^{\otimes 2} = \text{Sym.}(\mathscr{H} \otimes \mathscr{H})$ is the space spanned by $|\mathbf{p}_1, \mathbf{p}_2\rangle$ such that Bose statisitcs is followed and so on...

Lecture 12: Making things invariant

The states that we had defined earlier, $|\mathbf{p}_1, \mathbf{p}_2, \dots\rangle$ are *improper*, in the sense that they cannot be normalised. The obvious choice, $\langle \mathbf{p}|\mathbf{p}'\rangle = (2\pi)^3 {\delta^3}_{\mathbf{p}-\mathbf{p}'}$ is not physical due to the presence of the Dirac-delta. We can form physical states by 'smearing out' these deltas, by defining:

$$|\psi\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \psi(\mathbf{p}) |\mathbf{p}\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \psi(\mathbf{p}) \hat{a}^\dagger_{\mathbf{p}} ||\Omega\rangle\rangle$$

where $\psi(\mathbf{p})$ is an $\mathbb{L}^2$ function. However, this smearing process is not apriori Lorentz invariant. Thus, we can force the process to follow Lorentz invariance since the inception only, so that we do not run into trouble later.

12.1. Invariant Inner Product

For simplicity, consider a Lorentz boost along the x direction. Then,

$${p'}^1 = \gamma(p^1 + \beta E) \qquad\qquad E' = \gamma(E + \beta p^1)$$

where $\gamma = {}^1/\sqrt{1-({}^v/c)^2}$ and $\beta = {}^v/c$ are defined as usual. Now, note that $p^\mu p_\mu = m^2 \implies E^2 = (p^0)^2 + (p^1)^2 + (p^2)^2 + m^2 \implies 0 < E = \sqrt{(p^0)^2 + (p^1)^2 + (p^2)^2 + m^2}$ under the Minkowski metric. As the transformation is only along x direction, the momentum in x and y do not change. Thus, we have:

$$\frac{\partial E}{\partial p^j} = \frac{1}{2E} \times (2p^j) \implies \frac{\partial E}{\partial p^1} = \frac{p^1}{E}$$

Now, note the transformation of the Dirac-delta:

$$\int \mathrm{d}^3\mathbf{y}\ {\delta^3}_{\mathbf{x}-\mathbf{y}} = \int \mathrm{d}^3\mathbf{y}'\ {\delta^3}_{\mathbf{x}'-\mathbf{y}'} = \int \left|\frac{\partial \mathbf{y}'}{\partial \mathbf{y}}\right| \mathrm{d}^3\mathbf{y}\ {\delta^3}_{\mathbf{x}'-\mathbf{y}'}$$

Comparing, we see that the Dirac-delta transforms as ${\delta^3}_{\mathbf{x}-\mathbf{y}} = \left|\frac{\partial \mathbf{y}'}{\partial \mathbf{y}}\right| {\delta^3}_{\mathbf{x}'-\mathbf{y}'}$ where $\left|\frac{\partial \mathbf{y}'}{\partial \mathbf{y}}\right|$ is the determinant of the Jacobian matrix J of the transformation. Hence, for the Lorentz transformation, we have:

$$\begin{aligned}{\delta^3}_{\mathbf{p}-\mathbf{k}} &= \delta(p^1 - k^1)\delta(p^2 - k^2)\delta(p^3 - k^3) \\ &= \mathfrak{J}\ \delta(p'^1 - k'^1)\delta(p'^2 - k'^2)\delta(p'^3 - k'^3)\end{aligned}$$

Now, let us calculate the Jacobian matrix, whose elements will be given by $J^{ij} = \frac{\partial p'^i}{\partial p^j}$ (note that since we are concerned only with 3-momentum delta function, only spatial indices will appear). Note that $\frac{\partial p'^2}{\partial p^j} = \delta_{2j}$ and $\frac{\partial p'^3}{\partial p^j} = \delta_{3j}$. Only for $j = 1$, we have:

$$\frac{\partial p'^1}{\partial p^j} = \gamma \delta_{1j} + \gamma\beta \frac{\partial E}{\partial p^j}$$

and from the dispersion relation, we have $\frac{\partial E}{\partial p^j} = \frac{p^j}{E}$. The matrix is then:

$$J \equiv \begin{pmatrix} \gamma\left(1 + \beta\frac{p^1}{E}\right) & \gamma\beta\frac{p^2}{E} & \gamma\beta\frac{p^3}{E} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \implies \mathfrak{J} = \det J = \gamma\left(1 + \beta\frac{p^1}{E}\right) = \frac{\gamma}{E}\left(E + \beta p^1\right) = \frac{E'}{E}$$

Thus, we obtain the transformation of the delta function as:

$${\delta^3}_{\mathbf{p}-\mathbf{k}} = \frac{E'}{E}{\delta^3}_{\mathbf{p}'-\mathbf{k}'} \implies E{\delta^3}_{\mathbf{p}-\mathbf{k}} \text{ is invariant under L.T}$$

Then, we can now define the Lorentz-invariant inner product such that:

$$\langle \mathbf{p}|\mathbf{k}\rangle = (2E_{\mathbf{p}})(2\pi)^3{\delta^3}_{\mathbf{p}-\mathbf{k}} \equiv (2\omega_{\mathbf{p}})(2\pi)^3{\delta^3}_{\mathbf{p}-\mathbf{k}}$$

In line with this, we can therefore define a single-particle state as:

$$|\mathbf{p}\rangle = \sqrt{2\omega_{\mathbf{p}}}\hat{a}^\dagger_{\mathbf{p}}\,|\Omega\rangle$$

Technically, we should have made some distinct notation to differentiate between the states with and without proper normalisation, however, henceforth we will always use normalised states, so confusion might not be there.

12.2. Invariant Measure

Define the *projector* onto the single-particle state,

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

Here we have used the states with proper normalisation, obeying Lorentz invariance. The LHS is Lorentz invariant, since the operator essentially checks whether we have single particle or not, which is a frame independent fact. Thus, the only thing to change is the *integral measure*!

Consider the following integral,

$$\int \mathrm{d}^4 p\ \delta((p^0)^2 - \mathbf{p}^2 - m^2)\Theta(p^0)$$

Note that, $\mathrm{d}^4 p$ is Lorentz invariant. Also, from the dispersion relation, we have:

$$p^\mu p_\mu = (p^0)^2 - \mathbf{p}^2 = m^2$$

The dispersion relation is also Lorentz invariant, since it contains scalars $p^\mu p_\mu$ and m^2. Solving for p^0, we have $p^0 = \pm\sqrt{\mathbf{p}^2 + m^2}$. Now, the choice of which branch to choose (poisitive or negative) is also Lorentz invariant, since here we are considering only *orthochronous transformations*, which does not change the time component of the momentum. Thus, overall the integral is Lorentz-invariant.

$$\begin{aligned}
\int \mathrm{d}^4p\ \delta((p^0)^2 - \mathbf{p}^2 - m^2)\Theta(p^0) &= \int \mathrm{d}^4p\ \delta((p^0)^2 - E_{\mathbf{p}}^2)\Theta(p^0) \\
&= \int d^3\mathbf{p} \int_{-\infty}^{\infty} \mathrm{d}p^0\ \delta((p^0)^2 - E_{\mathbf{p}}^2)\Theta(p^0) \\
&= \int d^3\mathbf{p} \int_{-\infty}^{\infty} \mathrm{d}p^0\ \frac{1}{2E_{\mathbf{p}}}\big[\delta(p^0 + E_{\mathbf{p}}) + \delta(p^0 - E_{\mathbf{p}})\big]\Theta(p^0) \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{2E_{\mathbf{p}}}
\end{aligned}$$

Hence, whatever measure we were taking earlier, that is, $\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}$ was not Lorentz-invariant, however, as seen above,

$$\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} \text{ is manifestly Lorentz invariant}$$

Nice! So, henceforth we will always try to write the integrals in manifestly Lorentz invariant form, for example, the field expression becomes:

$$\hat{\phi}(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}\Big(\sqrt{2\omega_{\mathbf{p}}}\hat{a}_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + \sqrt{2\omega_{\mathbf{p}}}\hat{a}^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}}\Big)$$

We could technically absorb the $\sqrt{2\omega_{\mathbf{p}}}$ in the definition of the creation and annihilation operators which would make the new field expressions similar to the old ones, but it is not really needed!

Finally, we can rewrite the identity operator on one-particle states as,

$$\mathbb{1} := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}\left|\mathbf{p}\right\rangle\left\langle\mathbf{p}\right|$$

12.3. Position states?

Consider the field operator acting on the vacuum. Since annihilation operators kill the vacuum, we only have:

$$\hat{\phi}(x)\left|\Omega\right\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}e^{-i\mathbf{p}\cdot\mathbf{x}}\left|p\right\rangle$$

Apart from the factor $\frac{1}{2\omega_{\mathbf{p}}}$, this looks more or less like some position state since it is linear superposition of all well-defined momentum states. Also, note that:

$$\begin{aligned}
\left\langle\Omega\right|\hat{\phi}(\mathbf{x})\left|\mathbf{p}'\right\rangle &= \left\langle\Omega\right|\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}\Big(\sqrt{2\omega_{\mathbf{p}}}\hat{a}_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + \sqrt{2\omega_{\mathbf{p}}}\hat{a}^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}}\Big)\sqrt{2\omega_{\mathbf{p}}}\hat{a}^\dagger_{\mathbf{p}'}\left|\Omega\right\rangle \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(e^{i\mathbf{p}\cdot\mathbf{x}}\right)\left\langle\Omega\right|\hat{a}_{\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}'}\left|\Omega\right\rangle \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(e^{i\mathbf{p}\cdot\mathbf{x}}\right)\left\langle\Omega\right|(2\pi)^3\delta^3_{\mathbf{p}-\mathbf{p}'} + \hat{a}^\dagger_{\mathbf{p}'}\hat{a}_{\mathbf{p}}\left|\Omega\right\rangle \\
&= e^{i\mathbf{p}'\cdot\mathbf{x}}
\end{aligned}$$

We can interpret this as the position-space representation of the single-particle state $\left|\mathbf{p}\right\rangle$, exactly as in non-relativistic quantum mechanics. Thus, we have:

$$\left\langle\Omega\right|\hat{\phi}(\mathbf{x})\left|\mathbf{p}\right\rangle = \left\langle x|p\right\rangle = e^{i\mathbf{p}\cdot\mathbf{x}}$$

Lecture 13: Heisenberg Picture

Till now, we saw that the operator $\hat{\phi}(\mathbf{x})$ has no time dependence (we defined it at $t = 0$). However, space and time should be on equal footing. Hence, it is better to define these in the Heisenberg picture, where the operators evolve in time while the states don't. Recall that a Heisenberg operator is defined from the Schödinger operator via,

$$\mathcal{O}_H = e^{+iHt}\mathcal{O}_s e^{-iHt} \implies \frac{\mathrm{d}\mathcal{O}_H}{\mathrm{d}t} = i\,[H, \mathcal{O}_H] + \frac{\partial \mathcal{O}_H}{\partial t}$$

We consider here only free fields, that is, there is no interactions. However, in general, the Hamiltonian can be decomposed as

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int.}}$$

where $\hat{H}_0$ is the *free field*. In that case, it would be better to work in the interaction picture. However, for now we assume $\hat{H}_{\text{int.}} = 0$. Note that from Eq. 3 and Eq. 4, we can obtain the time evolution of the ladder operators as:

$$\frac{\mathrm{d}\hat{a}^\dagger_{\mathbf{p}}}{\mathrm{d}t} = i\omega_{\mathbf{p}}\hat{a}^\dagger_{\mathbf{p}} \implies \hat{a}^\dagger_{\mathbf{p}}(t) = \hat{a}^\dagger_{\mathbf{p}}(0)e^{i\omega_{\mathbf{p}}t} \implies e^{iH_0t}\hat{a}^\dagger_{\mathbf{p}}e^{-iH_0t} = \hat{a}^\dagger_{\mathbf{p}}e^{i\omega_{\mathbf{p}}t}$$
$$\frac{\mathrm{d}\hat{a}_{\mathbf{p}}}{\mathrm{d}t} = -i\omega_{\mathbf{p}}\hat{a}_{\mathbf{p}} \implies \hat{a}_{\mathbf{p}}(t) = \hat{a}_{\mathbf{p}}(0)e^{-i\omega_{\mathbf{p}}t} \implies e^{iH_0t}\hat{a}_{\mathbf{p}}e^{-iH_0t} = \hat{a}_{\mathbf{p}}e^{-i\omega_{\mathbf{p}}t}$$

Now, we calculate the field operator in the Heisenberg picture which gives:

$$\begin{aligned}
\hat{\phi}(\mathbf{x}, t) &= e^{iH_0t}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\Big(\hat{a}^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{a}_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}}\Big)e^{-iH_0t} \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\Big(e^{iH_0t}\hat{a}^\dagger_{\mathbf{p}}e^{-iH_0t}e^{-i\mathbf{p}\cdot\mathbf{x}} + e^{iH_0t}\hat{a}_{\mathbf{p}}e^{-iH_0t}e^{i\mathbf{p}\cdot\mathbf{x}}\Big) \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\Big(\hat{a}^\dagger_{\mathbf{p}}e^{i\omega_{\mathbf{p}}t}e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{a}_{\mathbf{p}}e^{-i\omega_{\mathbf{p}}t}e^{i\mathbf{p}\cdot\mathbf{x}}\Big) \\
&= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\Big(\hat{a}^\dagger_{\mathbf{p}}e^{ip\cdot x} + \hat{a}_{\mathbf{p}}e^{-ip\cdot x}\Big)
\end{aligned}$$

where $p \cdot x = p^\mu x_\mu = \omega_{\mathbf{p}}t - \mathbf{p}\cdot\mathbf{x}$ is the inner product of four momentum and four position in Minkowski space. Similarly, the conjugate momentum operator will be obtained as:

$$\hat{\pi}(\mathbf{x}, t) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}}\Big(\hat{a}_{\mathbf{p}}e^{-ip\cdot x} - \hat{a}^\dagger_{\mathbf{p}}e^{ip\cdot x}\Big)$$

Lecture 14 & 15: Propagators and Causality

In the previous section, we saw the time-evolution of the field operator in the Heisenberg pictre, essentially calculating an expression for $\hat{\phi}(\mathbf{x}, t)$. We will denote it simply as $\hat{\phi}(x)$ where x is the position four-vector while 3-vectors will be bold-faced.

Given a particle with momentum $|\mathbf{p}\rangle$, the transition to some other momentum state $|\mathbf{k}\rangle$, where $\mathbf{k} \neq \mathbf{p}$, is given by:

$$\langle\mathbf{k}|\mathbf{p}\rangle = (2\pi)^3\delta^3_{\mathbf{k}-\mathbf{p}} = 0$$

Since the momentum states are eigenstates of the Hamiltonian, transition is not possible. However, same cannot be said for position states. To understand if transition to another state is possible, we need something measurable. The obvious choice $\langle x|y\rangle = \langle\Omega|\,\hat{\phi}(x)\hat{\phi}(y)\,|\Omega\rangle$ is not always physical, since the field operators are not Hermitian in general (eg. complex scalar field). Nevertheless, let us see what happens when we calculate this quantity.

Formally, the inner product is called the *Wightman function* and in terms of the operators, it is defined as:

$$D(x-y) := \langle x|y\rangle = \langle\Omega|\,\hat{\phi}(x)\hat{\phi}(y)\,|\Omega\rangle$$

where x and y are four-positions and $\hat{\phi}$ is the field operator in the Heisenberg picture. The Wightman function is simply the 2-point correlation function. Let us calculate an explicit expression for this:

$$\begin{aligned} D(x-y) &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\frac{1}{\sqrt{2\omega_{\mathbf{p}'}}}\langle\Omega|\left\{\hat{a}_\mathbf{p}e^{-ip\cdot x}+\hat{a}^\dagger_\mathbf{p}e^{ip\cdot x}\right\}\left\{\hat{a}_{\mathbf{p}'}e^{-ip'\cdot y}+\hat{a}^\dagger_{\mathbf{p}'}e^{ip'\cdot y}\right\}|\Omega\rangle \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\frac{1}{\sqrt{2\omega_{\mathbf{p}'}}}e^{-ip\cdot x}e^{ip'\cdot y}\langle\Omega|\,\hat{a}_\mathbf{p}\hat{a}^\dagger_{\mathbf{p}'}\,|\Omega\rangle \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\frac{1}{\sqrt{2\omega_{\mathbf{p}'}}}e^{-ip\cdot x}e^{ip'\cdot y}(2\pi)^3\delta^3{}_{\mathbf{p}-\mathbf{p}'} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_\mathbf{p}}e^{-ip\cdot(x-y)} \end{aligned} \tag{5}$$

From the above expression, it is evident that $D(x-y)$ is Lorentz invariant since the measure and integrand both are invariant. Let us consider the case for different intervals.

Case 1: (x-y) is a *time-like* interval, that is, $(x-y)^2 > 0$ under Minkowski metric $(+,-,-,-)$

We can find a frame where the spatial interval is zero, that is, $\mathbf{x}'-\mathbf{y}'=0$. Now, since $x^0 > y^0$ in our original frame and we are considering *orthochronous* Lorentz transformation, $(x')^0 > (y')^0$ too, since orthochronous transformations cannot flip the sign of the time-component. Defining $t = (x')^0 - (y')^0$, we have:

$$\begin{aligned} D(x-y) = D(x'-y') &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_\mathbf{p}}e^{-ip^0 t} \\ &= \int_0^\infty \frac{4\pi p^2 dp}{(2\pi)^3}\frac{1}{2\sqrt{p^2+m^2}}e^{-i\sqrt{p^2+m^2}t} \\ &= \frac{1}{(2\pi)^2}\int_0^\infty dp\frac{p^2}{\sqrt{p^2+m^2}}e^{-i\sqrt{p^2+m^2}t} \end{aligned}$$

Let us take $\epsilon = \sqrt{p^2+m^2} \implies \epsilon^2 = p^2+m^2 \implies \epsilon\mathrm{d}\epsilon = p\mathrm{d}p$. Substituting this in the integral, we obtain:

$$\begin{aligned} D(x-y) &= \frac{1}{(2\pi)^2}\int_0^\infty dp\frac{p^2}{\sqrt{p^2+m^2}}e^{-i\sqrt{p^2+m^2}t} \\ &= \frac{1}{(2\pi)^2}\int_m^\infty \frac{1}{\not{\epsilon}}e^{-i\epsilon t}(\epsilon^2-m^2)\frac{\not{\epsilon}\mathrm{d}\epsilon}{\sqrt{\epsilon^2-m^2}} \\ &= \frac{1}{(2\pi)^2}\int_m^\infty e^{-i\epsilon t}\sqrt{\epsilon^2-m^2}\mathrm{d}\epsilon \end{aligned}$$

Now, let us take $\epsilon = m\cosh\theta \implies \mathrm{d}\epsilon = m\sinh\theta\mathrm{d}\theta$. Substituting this in the integral above, we obtain:

$$\begin{aligned}
D(x-y) &= \frac{1}{(2\pi)^2}\int_0^\infty e^{-im\cosh\theta t}\sinh^2\theta\mathrm{d}\theta \\
&= \frac{1}{2(2\pi)^2}\int_0^\infty e^{-im\cosh\theta t}(\cosh 2\theta - 1)\mathrm{d}\theta \\
&= \frac{1}{2(2\pi)^2}\left(\int_0^\infty e^{-im\cosh\theta t}\cosh 2\theta\mathrm{d}\theta - \int_0^\infty e^{-im\cosh\theta t}\mathrm{d}\theta\right) \\
&= \frac{1}{2(2\pi)^2}(K_2(imt) - K_0(imt))
\end{aligned}$$

where $K_\alpha(x) = \int_0^\infty e^{-x\cosh\theta}\cosh\alpha\theta\ \mathrm{d}\theta$ is the integral representation of the *modified Bessel function of the second kind.* The integral can also be solved in the asymptotic limit, when $t >> \frac{1}{m}$. Note that for large t and large ϵ, the phase in the exponent is large and there is rapid oscillations, leading to the overall integral being zero. Hence, the contribution to the integral will only come from around $\epsilon \approx m$.

Thus, taking $\epsilon = m + \xi, \xi << m$, we have:

$$D(x-y) = \frac{e^{-imt}}{(2\pi)^2}\int_0^\infty e^{-i\xi t}\sqrt{(2m+\cancel{\xi})\xi}\mathrm{d}\xi = \frac{e^{-imt}\sqrt{2m}}{(2\pi)^2}\int_0^\infty \sqrt{\xi}e^{-i\xi t}\mathrm{d}\xi = e^{-imt}\frac{\sqrt{2m}}{(2\pi)^2 t^{3/2}}\int_0^\infty \sqrt{z}e^{-iz}\mathrm{d}z$$

The integral is found out to be $\frac{\Gamma(3/2)}{i^{3/2}} = \frac{\sqrt{\pi}}{2}e^{-3\pi i/4}$, by the analytic continuation of the Gamma function. Anyways, what we found is that, upto some factor,

$$D(x-y) \sim e^{-imt}$$

Case 2: (x-y) is a *space-like* interval, that is, $(x-y)^2 < 0$ under the Minkowski metric $(+,-,-,-)$

We can find a frame where the temporal interval is zero, that is, $(x')^0 - (y')^0 = 0$. Let us define $\mathbf{r} = \mathbf{x} - \mathbf{y}$. Then we have:

$$D(x-y) = D(x'-y') = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{r}}$$

Consider $\mathbf{r}$ along the polar axis (if not, the rotate the polar axis), then we have:

$$\begin{aligned}
D(x-y) &= \frac{1}{2(2\pi)^2}\int_0^\infty \mathrm{d}p\int_{-1}^1 \mathrm{d}(\cos\theta)\ \frac{p^2}{\sqrt{p^2+m^2}}e^{ipr\cos\theta} \\
&= \frac{1}{2(2\pi)^2}\frac{1}{ir}\int_0^\infty \mathrm{d}p\frac{p}{\sqrt{p^2+m^2}}(e^{ipr} - e^{-ipr}) \\
&= \frac{1}{2ir(2\pi)^2}\int_{-\infty}^\infty \mathrm{d}p\frac{pe^{ipr}}{\sqrt{p^2+m^2}} \\
&= \frac{\mathcal{I}}{2ir(2\pi)^2}
\end{aligned}$$

Since square root is a multi-valued function, the integrand has branch points at $p = \pm im$ and we choose the branch cuts to be emanating from the poles and going to infinity. Let $r > 0$, so we chose the contour to close in the upper-half plane, as per Jordan's lemma. The contour, which avoids the branch cut, is shown below.

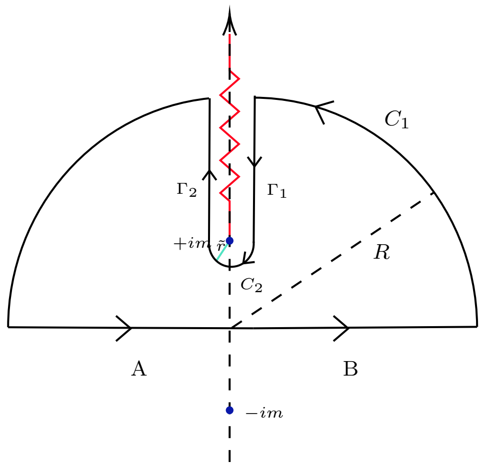
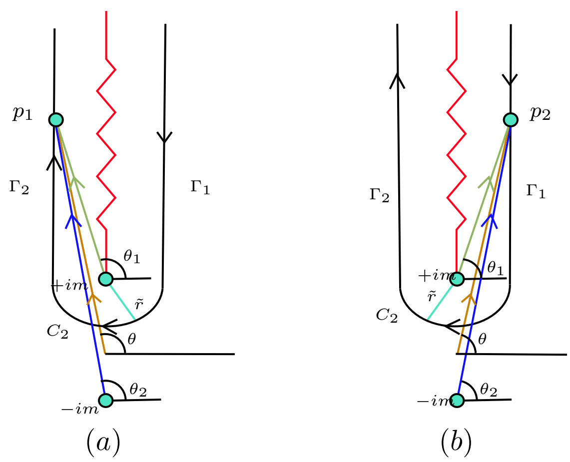


Figure 2: Contour for the integral (r>0). The blue dots represent the branch points while red kinked line denotes the branch cut.

The integrals along the arcs vanish for $R \to \infty$ and $\tilde{r} \to 0$ [1]. The only contribution to the integral comes because of the discontinuity across the branch cut. Since the contour encloses no poles, then by the Residue Theorem, we have:

$$\mathcal{I} + \int_{\Gamma_1} + \int_{\Gamma_2} = 0$$

Let us now see what these two integrals along both side of the branch cut evaluates to. For that, consider the diagram below:



Note that, on the left-side of the cut, θ_1 and θ_2 both are $\pi/2$. On the right side, since $p + im$ does not encircle the branch cut, θ_2 remain $\pi/2$ however, θ_1 becomes $\pi/2 - 2\pi = -3\pi/2$, since it encircles the branch cut. We write it as following:

$$\sqrt{p^2 + m^2} = \sqrt{|p + im||p - im|} \exp\left(i\frac{\theta_1 + \theta_2}{2}\right) = \begin{cases} \sqrt{|p + im||p - im|}i & \text{right} \\ -\sqrt{|p + im||p - im|}i & \text{left} \end{cases}$$

Then, the square root becomes negative of each other on either side of the branch cut (on right, side,).

[1] There is a subtlety tbh; for Jordan's lemma to be applicable, the integrand other than the exponential should vanish as $R \to \infty$. However, it doesn't happen here. Yet, by some transformation with Dirac delta, we can make this thing work afaik (Refer Padmanabhan, Problem1)!

Replacing $p = i\rho$, we have:

$$0 = \mathcal{I} + \int_{\infty}^{m} \frac{i\rho e^{-\rho r}}{\sqrt{\rho^2 - m^2}} + \int_{m}^{\infty} \frac{-i\rho e^{-\rho r}}{\sqrt{\rho^2 - m^2}}$$

$$\implies \mathcal{I} = 2i \int_{m}^{\infty} \frac{\rho e^{-\rho r}}{\sqrt{\rho^2 - m^2}}$$

Substituting $\rho = m\cosh\lambda$, we have the expression:

$$\mathcal{I} = 2im \int_{0}^{\infty} \cosh\lambda e^{-mr\cosh\lambda} = 2imK_1(mr)$$

Therefore, the final value of the correlation function becomes:

$$D(x-y) = \frac{m}{4\pi^2 r} K_1(mr) \xrightarrow{r\to\infty} \frac{m}{4\pi^2 r}\sqrt{\frac{\pi}{2mr}} e^{-mr} \sim e^{-mr}$$

We see that for spacelike separations, the amplitude falls off as $e^{-mr} \neq 0$. This shows that there is a non-zero probability amplitude for a particle to propagate over a spacelike interval, which violates causality (as it implies faster than light travel). As we had said earlier, this quantity is not very physical. A more appropriate quantity to analyse will be the commutator of the fields, that is,

$$\begin{aligned}\left[\hat{\phi}(x), \hat{\phi}(y)\right] &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{\mathrm{d}^3\mathbf{p'}}{(2\pi)^3} \frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p'}}}} \left\{\left[\hat{a}_{\mathbf{p}}, \hat{a}^\dagger_{\mathbf{p'}}\right] e^{-ip\cdot x} e^{ip'\cdot y} + \left[\hat{a}^\dagger_{\mathbf{p'}}, \hat{a}_{\mathbf{p'}}\right] e^{ip\cdot x} e^{-ip'\cdot y}\right\} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right) \\ &= D(x-y) - D(y-x)\end{aligned}$$

Thus, we define the propagator as:

$$\triangle(x-y) = D(x-y) - D(y-x)$$

In this case, for a space-like interval, we can find a transformation which will take $(x-y) \mapsto (y-x)$, that is, If a particle can travel in a spacelike direction from $x \to y$, it can just as easily travel from $y \to x$ (since there is not Lorentz invariant way to order events), as a result of which, $\triangle(x-y) = 0$, no longer causing any violation of causality. For time-like separation, we cannot find such a transformation and as a result, the commutator is non-zero.

15.1. Propagators

Note that the commutator is basically a number (with an implicit identity operator sitting quietly) and thus,

$$\langle\Omega| \left[\hat{\phi}(x), \hat{\phi}(y)\right] |\Omega\rangle = D(x-y) - D(y-x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)$$

Let us now calculate a 4-D integral of the following form, taking suppose $x_0 > y_0$:

$$\begin{aligned}\mathcal{I} &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{\mathrm{d}p^0}{2\pi i} \frac{-e^{-ip\cdot(x-y)}}{p^2 - m^2} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \int \frac{\mathrm{d}p^0}{2\pi i} \frac{-e^{-ip^0(x^0-y^0) + i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{(p^0)^2 - |\mathbf{p}|^2 - m^2} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \int \frac{\mathrm{d}p^0}{2\pi i} \frac{-e^{-ip^0(x^0-y^0) + i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{(p^0)^2 - \omega_{\mathbf{p}}^2} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \int \frac{\mathrm{d}p^0}{2\pi i} \frac{-e^{-ip^0(x^0-y^0)}}{[(p^0) - \omega_{\mathbf{p}}][(p^0) + \omega_{\mathbf{p}}]}\end{aligned}$$

Note that the integral has two poles at $p^0 = \pm\omega_{\mathbf{p}}$, which lies on the real line. Since we are having $x^0 > y^0$, to apply Jordan's lemma, we need to close the contour in the lower half-plane. For this case, we choose to include both the poles into the contour, hence, applying the Residue Theorem, we have:

$$\int \mathrm{d}p^0 \frac{e^{-ip^0(x^0-y^0)}}{[(p^0)-\omega_{\mathbf{p}}][(p^0)+\omega_{\mathbf{p}}]} = -2\pi i\left(\frac{e^{-i\omega_{\mathbf{p}}(x^0-y^0)}}{2\omega_{\mathbf{p}}} + \frac{e^{i\omega_{\mathbf{p}}(x^0-y^0)}}{-2\omega_{\mathbf{p}}}\right)$$

$-2\pi i$ comes since the contour is taken clockwise. Substituting this value in the above integral, we obtain:

$$\begin{aligned}\mathcal{I} &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\left(\frac{e^{-i\omega_{\mathbf{p}}(x^0-y^0)}}{2\omega_{\mathbf{p}}} + \frac{e^{i\omega_{\mathbf{p}}(x^0-y^0)}}{-2\omega_{\mathbf{p}}}\right)\\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(\frac{e^{-i\omega_{\mathbf{p}}(x^0-y^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{2\omega_{\mathbf{p}}} + \frac{e^{i\omega_{\mathbf{p}}(x^0-y^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{-2\omega_{\mathbf{p}}}\right)\\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}\left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)\end{aligned}$$

In the second term, we had taken $\mathbf{p} \to -\mathbf{p}$. Since $\omega_{\mathbf{p}} = \omega_{-\mathbf{p}}$ and the volume element remains invariant, we obtained the second term in the final expression.

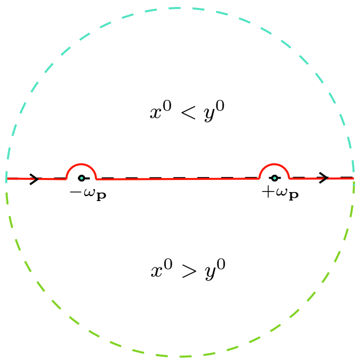


Figure 3: The contour choice for the *retarded propagator*. The green contour gives non-zero value while the blue contour gives zero since no pole is inside the contour. The read curve is common to both the contours

Thus, we see from the final expression that,

$$\langle\Omega|\left[\hat{\phi}(x),\hat{\phi}(y)\right]|\Omega\rangle = \int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\int\frac{\mathrm{d}p^0}{-2\pi i}\frac{e^{-ip\cdot(x-y)}}{(p^0)^2-|\mathbf{p}|^2-m^2}$$

Note that, instead of deforming the contour, we could have just shifted the poles downwards (by some amount $i\epsilon$ where $\epsilon > 0$) from the real axis, so that they fall into the contour which belongs completely to the lower half-plane. Thus, we can write the expression equivalently as:

$$\langle\Omega|\left[\hat{\phi}(x),\hat{\phi}(y)\right]|\Omega\rangle = \int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\int\frac{\mathrm{d}p^0}{-2\pi i}\frac{e^{-ip\cdot(x-y)}}{(p^0+i\epsilon)^2-|\mathbf{p}|^2-m^2} = \int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\int\frac{\mathrm{d}p^0}{-2\pi i}\frac{e^{-ip\cdot(x-y)}}{(p^0+i\epsilon)^2-\omega_{\mathbf{p}}^2}$$

Note that, if we had taken $x^0 < y^0$, we would have closed the contour in the upper half-plane for Jordan's lemma to be valid and we would have got the integral as zero, since the contour did not enclose any of the poles. This is called the *Retarded Propagator* and is written compactly as,

$$D_R(x-y) = \Theta(x^0-y^0)\langle\Omega|\left[\hat{\phi}(x),\hat{\phi}(y)\right]|\Omega\rangle$$

where $\Theta(\cdot)$ represents the Heaviside-theta function. Similar to the above, we can define the *Advanced Propagator* which vanishes for $x^0 > y^0$. In this case, the poles are shifted upwards by $i\epsilon$. We can define this as:

$$D_A(x-y) = \Theta(y^0-x^0)\langle\Omega|\left[\hat{\phi}(x),\hat{\phi}(y)\right]|\Omega\rangle$$

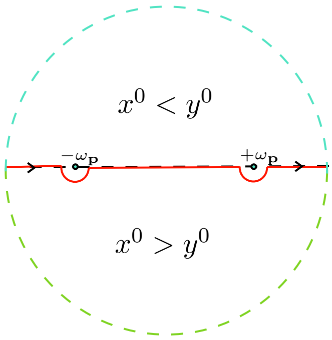


Figure 4: Contour choice for the advanced propagator.

There exists another pole prescription which gives rise to the celebrated *Feynman Propagator* which we investigate now.

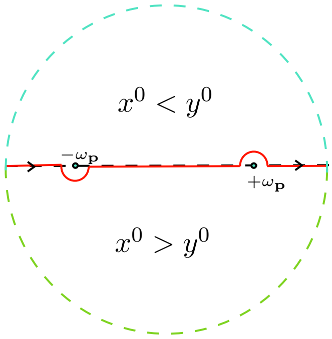


Figure 5: Pole choice for the Feynman propagator. We choose one pole to be inside and one pole to be outside the contour.

The Feynman propagator is defined as:

$$D_F(x-y) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \int \frac{\mathrm{d}p^0}{-2\pi i} \frac{e^{-ip\cdot(x-y)}}{(p^0)^2 - \omega_{\mathbf{p}}^2 + i\epsilon}$$

In this case, the contour is chosen such that one pole is shifted up and one pole is shifted down. Let the poles be,

$$p^0 = -\omega_{\mathbf{p}} + i\epsilon' \qquad p^0 = \omega_{\mathbf{p}} - i\epsilon'$$

Then we have the following,

$$(p^0 + \omega_{\mathbf{p}} - i\epsilon')(p^0 - \omega_{\mathbf{p}} + i\epsilon') = (p^0)^2 - (\omega_{\mathbf{p}} - i\epsilon')^2 = (p^0)^2 - \omega_{\mathbf{p}}^2 + \epsilon'^2 + 2i\omega_{\mathbf{p}}\epsilon'$$

We can neglect the ϵ'^2 term compared to the other term since it's a infinitesimal quantity. Also, for the same reason, we can define $\epsilon \equiv 2\omega_{\mathbf{p}}\epsilon'$ in the integral. The Feynman propagator can then be written as,

$$\begin{aligned} D_F(x-y) &= \Theta(x^0 - y^0) \langle\Omega| \hat{\phi}(x)\hat{\phi}(y) |\Omega\rangle + \Theta(y^0 - x^0) \langle\Omega| \hat{\phi}(y)\hat{\phi}(x) |\Omega\rangle \\ &= \langle\Omega| \mathcal{T}\hat{\phi}(x)\hat{\phi}(y) |\Omega\rangle \end{aligned}$$

where $\mathcal{T}$ is the *time-ordering* operator, which brings the operator to the left with the 'greater' time, similar to the normal ordering operator encounteres previously.

Feynman propagator is very important and has many applications as will be discussed later.

Lecture 16: Introducing the Dirac Equation

We saw that, when quantised, the Klein-Gordon equation provided a negative probability density, which was problematic since it had no obvious interpretation. Hence we said that, the KG equation cannot describe a single-particle system, instead, it describes a spin-0 particle *field*. All the problems arose since, from the relativistic dispersion relation,

$$E^2 = p^2 + m^2 = p_x^2 + p_y^2 + p_z^2 + m^2 \tag{6}$$

we obtain a differential equation which is second order in both space and time.

Let us now look for something which we force to be first order in space and time from the beginning. The most sane choice is to just take the square-root of the dispersion relation, however, square-root of an operator is ill-defined and has many other subtleties. We want everything to be as simple as possible. Hence, we try to *linearise* the square-root. We thus write,

$$E = \sqrt{p^2 + m^2} = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m$$

We first promote the momentum and energy to operators, that is, $p \to -i\nabla$ and $E \to i\partial_t$. Then the Schrödinger equation becomes,

$$i\partial_t \psi = \left(-i\alpha_i \frac{\partial}{\partial x^i} + \beta m\right)\psi \qquad \Longleftrightarrow \qquad H\psi = E\psi$$

Since we want this to match with the actual dispersion relation Eq. 6, let us see what happens when we apply the Hamiltonian twice on the time-independent equation (which would give us an E^2).

Note that, since the *linearisation* of the square-root is extremely non-trivial, α_i and β might not be just numbers, they could be literally anything. Thus, we do not make any apriori assumptions![1]

$$H^2 \equiv = \left(-\alpha_i^2 \frac{\partial^2}{\partial {x^i}^2} + \beta m^2 - (\alpha_i\alpha_j + \alpha_j\alpha_i)\frac{\partial^2}{\partial x^i \partial x^j} - im(\beta a_i + \alpha_i \beta)\frac{\partial}{\partial x^i}\right)$$

If the relativistic dispersion relation has to be satisfied, we find some constraints on α_i and β:

$$\begin{aligned} \alpha_i^2 &= \mathbb{1} \qquad & \alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij}\mathbb{1} \\ \beta^2 &= \mathbb{1} \qquad & \beta\alpha_i + \alpha_i\beta &= 0 \end{aligned}$$

where $\mathbb{1}$ represents the identity operator in appropriate space. Having α_i and β to be just simple scalar numbers clearly do not satisfy the equations, hence let us assume for now, that these are matrices. We then additionally impose the contraint that $\alpha_i^\dagger = \alpha_i$ and $\beta^\dagger = \beta$ since energy and momentum operator (and hence the Hamiltonian) are Hermitian. With this assumption, let us note some nice properties of these matrices.

- Let v be an eigenvector of $\alpha_i \implies \alpha_i v = \lambda v \implies \underbrace{\alpha_i^2}_{\mathbb{1}} v = \lambda^2 v \implies \lambda = \pm 1$. Similar argument goes for β. Thus, these matrices have eigenvalues either +1 or -1.
- $\alpha_i\beta = -\beta\alpha_i \implies \alpha_i\beta^2 = -\beta\alpha_i\beta \implies \alpha_i = \beta\alpha_i\beta \implies \operatorname{Tr}\alpha_i = -\operatorname{Tr}\alpha_i \implies \operatorname{Tr}\alpha_i = 0$, where we have used the cyclic property of trace. Similar argument goes for β. Hence we obtain that these matrices are also traceless.
- Since the trace, that is, the sum of the eigenvalues, should equal zero, we should definitely have half of the eigenvalues as +1 and the other half to be -1. Thus, these matrices are guaranteed to be *even* dimensional.
- The relations between α_i looks suspiciously similar to the Pauli matrices and we can choose α_i to be σ_i. Alas, we run into a dangerous problem. We note that β cannot be the a two-dimensional matrix, since then it would not be consistent with the anticommutation relation with $\alpha_i \equiv \sigma_i$. Thus, the next option is to try for *four dimensional* matrices.

[1] Well, we do assume that α_i and $\frac{\partial}{\partial x^i}$ commute in some sense

Let us define the following,

$$\alpha_i := \left[\begin{array}{c:c} 0 & \sigma_i \\ \hdashline \sigma_i & 0 \end{array}\right] \qquad \beta := \left[\begin{array}{c:c} \mathbb{1} & 0 \\ \hdashline 0 & -\mathbb{1} \end{array}\right]$$

With this definition, we can indeed verify that all the conditions hold [1] and we obtain a way to proceed further. We have an expression for the equation as,

$$i\frac{\partial \psi}{\partial t} = -i\alpha \cdot \nabla + \beta m \implies i\beta\frac{\partial \psi}{\partial t} = -i\beta\alpha \cdot \nabla + \underbrace{\beta^2}_{\mathbb{1}} m$$

We would like to write the Dirac equation in covariant notation. To do so, let us introduce the following matrices:

$$\gamma^0 := \beta$$
$$\gamma^i := \gamma^0 \alpha_i$$

These matrices are called *Dirac matrices.* Then, using this definition in the above expression, we have:

$$i\gamma^0\partial_0\psi + i\gamma^i\partial_i\psi - m\psi = 0 \implies (i\gamma^\mu\partial_\mu - m)\psi = 0$$

Note that, we had introduced the 4-vector concept in an ad-hoc manner. However, doing this, we finally derive the well-known covariant form of the *Dirac equation*!

16.1. Form of Dirac Matrices

The Dirac matrices that we had defined have the following form:

$$\gamma^i = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}\begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{7}$$

Lecture 17: Dirac in Fourier Mode

We saw that in the previous section that the Dirac equation is essentially:

$$i\partial_t\psi = (-i\alpha \cdot \nabla + \beta m)\psi \qquad \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}$$

The solution is a four-component vector since the matrices are 4×4 dimensional. [2] Let the proposed solution be a plane-wave solution, since there is no interaction, and we decompose it into its Fourier modes, that is,

$$\psi(x) = e^{-iEt}\psi(\mathbf{x}) = e^{-iEt}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\psi_p e^{-i\mathbf{p}\cdot\mathbf{x}}$$

Substituting this in the equation above, we have:

- LHS:
$$i\partial_t\psi = i(-iE)\psi(x) = E\psi(x)$$

- RHS:
$$e^{-iEt}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(-i\alpha \cdot (i\mathbf{p}) + \beta m)\psi_p e^{i\mathbf{p}\cdot\mathbf{x}} = e^{-iEt}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(\alpha \cdot \mathbf{p} + \beta m)\psi_p e^{i\mathbf{p}\cdot\mathbf{x}}$$

[1] Note that these matrices are 4×4 matrices. Each element in the matrix are indeed 2×2 matrices.
[2] In non-relativistic QM, during angular momentum, we saw that multi-component wavefunctions imply a non-trivial spin angular momentum, so we expect Dirac equation to describe spinful particles.

Equating both LHS and RHS we get an equation in ψ_p,

$$\begin{aligned} 0 &= (\alpha\cdot\mathbf{p}+\beta m - E)\psi_p \\ &= \left(\begin{pmatrix} 0 & \sigma_i p_i \\ \sigma_i p_i & 0\end{pmatrix} + \begin{pmatrix}(m-E)\mathbb{1} & 0 \\ 0 & -(E+m)\mathbb{1}\end{pmatrix}\right)\psi_p \\ &= \begin{pmatrix}(m-E)\mathbb{1} & \sigma\cdot\mathbf{p} \\ \sigma\cdot\mathbf{p} & -(m+E)\mathbb{1}\end{pmatrix}\psi_p \end{aligned}$$

Now, for a solution to exist, the matrix should be *singular*, that is,

$$\det\left(\begin{pmatrix}(m-E)\mathbb{1} & \sigma\cdot\mathbf{p} \\ \sigma\cdot\mathbf{p} & -(m+E)\mathbb{1}\end{pmatrix}\right) = 0$$

We can calculate the determinant in the "usual" sense for block-matrices also, if the elements in the first column commute. Since first element is the identity matrix, it commutes with $\sigma\cdot\mathbf{p}$ and hence, we can write:

$$\begin{aligned} 0 &= \det\left((E^2-m^2)\mathbb{1} - (\sigma\cdot\mathbf{p})(\sigma\cdot\mathbf{p})\right) \\ &= \det\left((E^2-m^2)\mathbb{1} - (\sigma_i p_i)(\sigma_j p_j)\right) \\ &= \det\left((E^2-m^2)\mathbb{1} - (\sigma_i\sigma_j)p_i p_j\right) \\ &= \det\left((E^2-m^2)\mathbb{1} - |\mathbf{p}|^2 + \sum_{i<j}(\sigma_i\sigma_j+\sigma_j\sigma_i)p_i p_j\right) \\ &= \det\left((E^2-m^2-|\mathbf{p}|^2)\mathbb{1}\right) \end{aligned} \tag{8}$$

From this, we finally obtain that $E^2-m^2-|\mathbf{p}|^2 = 0 \implies E = \pm\sqrt{m^2+|\mathbf{p}|^2}$. We can see that, the equation says that for each momentum $\mathbf{p}$, we have two positive-energy solutions and two negative-energy solutions (each energy is *doubly degenerate*). Now let us find the eigenvectors. Assume it to be of the form,

$$\psi_p = \begin{pmatrix}\phi \\ \xi\end{pmatrix}$$

Then, substituting this in the eigenvalue equation we obtain two equations:

$$\begin{aligned} 0 &= (m-E)\phi + (\sigma\cdot\mathbf{p})\xi \\ 0 &= (\sigma\cdot\mathbf{p})\phi - (m+E)\xi \end{aligned}$$

Solving for ϕ and ξ, we get $\xi = \frac{(\sigma\cdot\mathbf{p})\phi}{m+E}$ and $\phi = \frac{(\sigma\cdot\mathbf{p})\xi}{E-m}$. Thus, we can write the solutions as:

$$\psi_p = \begin{pmatrix}\phi \\ \frac{(\sigma\cdot\mathbf{p})}{m+E}\,\phi\end{pmatrix} \qquad \text{and} \qquad \psi_p = \begin{pmatrix}\frac{(\sigma\cdot\mathbf{p})}{E-m}\,\xi \\ \xi\end{pmatrix} \tag{9}$$

Now, let us first take the particle at rest, that is, $\mathbf{p}=0$. Then the eigenvalues are just $\pm m$ and the equation reduces to:

$$\begin{pmatrix} m & 0 \\ 0 & -m\end{pmatrix}\begin{pmatrix}\phi \\ \xi\end{pmatrix} = \pm m\begin{pmatrix}\phi \\ \xi\end{pmatrix}$$

Since the matrix is diagonal, the eigenvectors are trivially,

$$(\psi_p)_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \quad (\psi_p)_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \quad (\psi_p)_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \quad (\psi_p)_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

We can check that, the first two correspond to the positive energy solutions and the other two correspond to the negative energy solutions. Now, let us come back to $\mathbf{p}\neq 0$.The matrix $\sigma\cdot\mathbf{p}$ comes in the expression quite often and calculating that, we find that:

$$\sigma\cdot\mathbf{p} \equiv \begin{pmatrix} p_3 & p_1 - ip_2 \\ p_1 + ip_2 & -p_3\end{pmatrix}$$

Now, let us take a vector v_m whose m^{th} position has 1 and all other elements are zero.

$$(Mv_m)_i = M_{ij}(v_m)j = M_{ij}\delta_{mj} = M_{im}$$

- CASE 1: $\phi = \begin{pmatrix}1\\0\end{pmatrix} \implies \xi = \frac{1}{E+m}\begin{pmatrix}p_3\\p_1+ip_2\end{pmatrix}$
- CASE 2: $\phi = \begin{pmatrix}0\\1\end{pmatrix} \implies \xi = \frac{1}{E+m}\begin{pmatrix}p_1-ip_2\\-p_3\end{pmatrix}$
- CASE 3: $\xi = \begin{pmatrix}1\\0\end{pmatrix} \implies \phi = \frac{1}{E-m}\begin{pmatrix}p_3\\p_1+ip_2\end{pmatrix}$
- CASE 4: $\xi = \begin{pmatrix}0\\1\end{pmatrix} \implies \phi = \frac{1}{E-m}\begin{pmatrix}p_1-ip_2\\-p_3\end{pmatrix}$

Thus, the final solutions for non-zero momentum are:

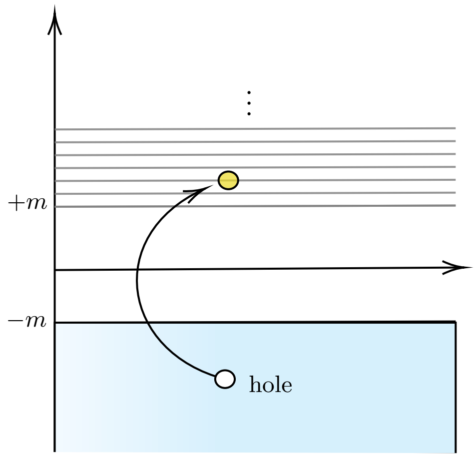
$$(\psi_p)_1 = \mathcal{N}_1\begin{pmatrix}1\\0\\\dfrac{p_3}{E+m}\\\dfrac{p_1+ip_2}{E+m}\end{pmatrix} \quad (\psi_p)_2 = \mathcal{N}_2\begin{pmatrix}0\\1\\\dfrac{p_1-ip_2}{E+m}\\\dfrac{-p_3}{E+m}\end{pmatrix} \quad (\psi_p)_3 = \mathcal{N}_3\begin{pmatrix}\dfrac{p_3}{E-m}\\\dfrac{p_1+ip_2}{E+m}\\1\\0\end{pmatrix} \quad (\psi_p)_4 = \mathcal{N}_4\begin{pmatrix}\dfrac{p_1-ip_2}{E-m}\\\dfrac{-p_3}{E+m}\\0\\1\end{pmatrix}$$

where $\mathcal{N}_i$ are some normalisation constants. The four components wavefunctions are called *Dirac spinors*.

Note that, if $\mathbf{p} = 0$ then the first two solutions correspond to the positive energy while the other two correspond to the negative energy. And thus, we define these solutions corresponding the $E = \sqrt{m^2 + |\mathbf{p}|^2}$ and $E = -\sqrt{m^2 + |\mathbf{p}|^2}$ respectively.

Note that the positive and negative energy solutions are separated by a gap of $2m$ as $E_+ = \sqrt{|\mathbf{p}|^2 + m^2} > m$ and $E_- = -\sqrt{|\mathbf{p}|^2 + m^2} < -m$. Also, the energy spectrum is *unbounded* from below.

It is an evident difficulty since any particle satisfying Dirac equation, will inevitably fall down to the lower, negative energy states, causing *destabilisation*.



To sort out this problem, the initial interpretation of the negative energy solutions were that the ground state (vacuum) was such that all the negative energy states were fully filled (if it were not, then transition to lower energy would always have been a possibility!). Thus, the *ground state* contains an infinite number of particles in the negative energy states.

The Dirac equation describes particles with spin-$\frac{1}{2}$ and hence a positive energy particle cannot transit to negative energy state due to the exclusion principle.

Excitations can happen such that a particle with negative energy occupy a positive energy state, leaving a 'hole' in the negative energy domain, which represented an *anti-particle*, having the same mass and *internal quantum numbers*.

Lecture 18: Few Properties of Dirac Equation

We will now look at some properties of the Dirac equation and also of the Dirac matrices. The first to check is the Lorentz covariance, (that is, under a Lorentz transformation, the equation transforms in a similar way), since we started with the relativistic dispersion relation.

18.1. Lorentz Covariance

Let us consider a Lorentz transformation,

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

Let us assume that under the transformation, the wavefunction changes as,

$$\psi'(x') = \psi'(\Lambda x) = S(\Lambda)\psi(x)$$

$S(\Lambda)$ is a *spinor* representation of the Lorentz transformation. What we assume here is that the Lorentz transformation on the wavefunction can be faithfully represented by the matrix S which depends only on the parameter of the transformation Λ and not on the spacetime. Also, since the transformation matrices are *invertible*, we have:

$$\psi(x) = S^{-1}(\Lambda)\psi'(x')$$

$$\begin{aligned}
0 &= (i\gamma^\mu\partial_\mu - m)\psi(x) \\
&= (i\gamma^\mu\Lambda^\nu{}_\mu\partial'_\mu - m)S^{-1}(\Lambda)\psi'(x') \\
&= (iS(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu\partial'_\mu - mS(\Lambda)S^{-1}(\Lambda))\psi'(x') \\
&= (iS(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu\partial'_\mu - m)\psi'(x')
\end{aligned}$$

Note that we had used that the matrix S is spacetime independent and thus, commutes with ∂_μ operator. The above equation is Lorentz covariant if we can find an S such that,

$$[S(\Lambda)\gamma^\mu S^{-1}(\Lambda)]\Lambda^\nu{}_\mu = \gamma^\nu \implies S^{-1}(\Lambda)S(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu S(\Lambda) = S^{-1}(\Lambda)\gamma^\nu S(\Lambda) \implies \Lambda^\nu{}_\mu\gamma^\mu = S^{-1}(\Lambda)\gamma^\nu S(\Lambda) \tag{10}$$

Thus, if we define $\gamma'^\mu \equiv \Lambda^\nu{}_\mu\gamma^\mu = S^{-1}(\Lambda)\gamma^\nu S(\Lambda)$, then we can say that the Dirac matrices transform with one copy of Λ and hence, the vector index on γ, which was introduced in an ad-hoc manner, can actually be taken seriously.

18.2. Helicity

Let us define the following matrices:

$$\Sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu,\gamma^\nu] \qquad \Sigma^i = \frac{1}{2}\epsilon^{ijk}\Sigma^{jk}$$

Then, we can calculate the form of the matrix given by Σ^i since,

$$\begin{aligned}
-2i\Sigma^{jk} &= \gamma^j\gamma^k - \gamma^k\gamma^j \\
&= \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \\
&= -\begin{pmatrix} \sigma_j\sigma_k & 0 \\ 0 & \sigma_j\sigma_k \end{pmatrix} + \begin{pmatrix} \sigma_k\sigma_j & 0 \\ 0 & \sigma_k\sigma_j \end{pmatrix} \\
&= -\begin{pmatrix} [\sigma_j,\sigma_k] & 0 \\ 0 & [\sigma_j,\sigma_k] \end{pmatrix} \\
&= -2i\epsilon^{jkl}\begin{pmatrix} \sigma_l & 0 \\ 0 & \sigma_l \end{pmatrix}
\end{aligned}$$

From this, we obtain a neat expression that:

$$\Sigma^{jk} = \epsilon^{jkl}\begin{pmatrix}\sigma_l & 0 \\ 0 & \sigma_l\end{pmatrix}$$

Next we use the expression given above:

$$\Sigma^i = \frac{1}{2}\epsilon^{ijk}\epsilon^{jkl}\begin{pmatrix}\sigma_l & 0 \\ 0 & \sigma_l\end{pmatrix} = \delta_{il}\begin{pmatrix}\sigma_l & 0 \\ 0 & \sigma_l\end{pmatrix} = \begin{pmatrix}\sigma_i & 0 \\ 0 & \sigma_i\end{pmatrix}$$

Let us calculate the commutator between the Hamiltonian and the quantity $\mathbf{\Sigma}\cdot\mathbf{p}$

$$\begin{aligned}[H, \mathbf{\Sigma}\cdot\mathbf{p}] &\sim [\alpha\cdot\mathbf{p} + \beta m, \mathbf{\Sigma}\cdot\mathbf{p}] \\ &= [\alpha\cdot\mathbf{p}, \mathbf{\Sigma}\cdot\mathbf{p}] + m\,[\beta, \mathbf{\Sigma}\cdot\mathbf{p}] \\ &= \begin{pmatrix}0 & \sigma\cdot\mathbf{p} \\ \sigma\cdot\mathbf{p} & 0\end{pmatrix}\begin{pmatrix}\sigma\cdot\mathbf{p} & 0 \\ 0 & \sigma\cdot\mathbf{p}\end{pmatrix} - \begin{pmatrix}\sigma\cdot\mathbf{p} & 0 \\ 0 & \sigma\cdot\mathbf{p}\end{pmatrix}\begin{pmatrix}0 & \sigma\cdot\mathbf{p} \\ \sigma\cdot\mathbf{p} & 0\end{pmatrix} + 0 \\ &= 0\end{aligned}$$

We thus see that this quantity indeed commutes with the Hamiltonian and hence we can measure the value of this operator simultaneously with the Hamiltonian. To provide a more physical measure, we define the *helicity* operator as:

$$\hat{h} := \frac{\hbar}{2}\frac{\mathbf{\Sigma}\cdot\mathbf{p}}{|\mathbf{p}|} \tag{11}$$

Supose our momentum is along the z-direction, then

$$\hat{h} = \frac{\hbar}{2}\Sigma^3 = \frac{\hbar}{2}\begin{pmatrix}1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1\end{pmatrix}$$

which has eigenvalues ± 1. Then, applying this operator on each of the four solutions, we will get eigenvalues ± 1 for positive and ± 1 for the negative solutions and thus the degeneracy is lifted.

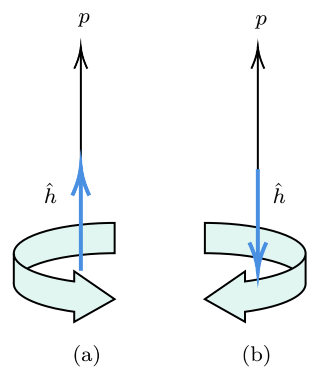


Figure 6: Fig (a) describes when $\hat{h}$ is along the momentum direction, giving positive helicity or right-handedness while Fig (b) describes when $\hat{h}$ is anti-parallel to the momenutm direction, giving negative helicity or left-handedness.

It can be shown that helicity is not Lorentz invariant and this is problematic. Only for massless particles, helicity becomes Lorentz invariant.

18.3. Properties of Dirac Matrices

We first note that,

$$(\gamma^0)^2 = \mathbb{1} \qquad (\gamma^i)^2 = -\mathbb{1}$$

Proof. We have $\gamma^0\gamma^0 = \beta^2 = \mathbb{1}$ and $(\gamma^i)^2 = \beta\alpha_i\beta\alpha_i = -\alpha_i\beta^2\alpha_i = -\alpha_i^2 = -\mathbb{1}$

This is a nice property with which we can differentiate between the spatial and temporal indices. Indeed, the Dirac matrices are somewhat used in relativity too, to describe spacetime.

It turns out that any 4×4 matrix can be written in terms of some combinations of Dirac matrices. In other words, the Dirac matrices form a basis of four dimensional square matrices. Note that, the obvious choice $\mathbb{1}$ and γ^μ cannot generate all the 4×4 matrices since we need 16 elements in the basis. The other elements in the basis are formed by multiplying the Dirac matrices:

- Two at a time, $\gamma^\mu \gamma^\nu$ ...... $\binom{4}{2} = 6$ matrices
- Three at a time, $\gamma^\mu \gamma^\nu \gamma^\alpha$ ...... $\binom{4}{3} = 4$ matrices
- Four at a time, $\gamma^\mu \gamma^\nu \gamma^\alpha \gamma^\beta$ ...... $\binom{4}{4} = 1$ matrix

Thus, we have a total of $1 + 4 + 6 + 4 + 1 = 16$ elements for the basis and we are happy! If we define,

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$$

Then the basis set can be written in terms of the 'fifth' Dirac matrix:

$$\Gamma = \{\mathbb{1}, \gamma^\mu, \Sigma^{\mu\nu}, \gamma^5, \gamma^5\gamma^\mu\}$$

Well, these matrices form what is called the *Clifford Algebra*, a fairly complicated mathematical concept. However, the crux of that is the *anti-commutation* relations satisfied by the Dirac matrices. Like, γ^μ's form the Clifford algebra in four dimensions while γ^μ along with γ^5 forms the Clifford algebra in five dimensions. There are higher-dimensional generalisations of the gamma matrices for this purpose.

Having claimed that the Dirac matrices form a basis of 4×4 matrices, have to prove that it is indeed a basis. For this, we need to show the linear independence of each basis element. Also, it would be nice if these basis elements somehow turn out to be 'orthogonal', since it is always convenient to work with orthonormal basis.

For the space of matrices, the most obvious choice to define orthognality is through the trace. That is, matrices A and B are orthogonal if $\mathrm{Tr}(AB) = 0$ [1]. Thus, it becomes necessary to work out the traces of the Dirac matrices.

Lecture 19: Trace and Contraction Identities

In the previous section, we saw that calculating the trace of the Dirac matrices become really important since these define the inner products. Let us look at some of the trace identities followed by the Dirac matrices.

19.1. Trace Identitites

- $\mathrm{Tr}(\gamma^\mu) = 0$

  Proof. We defined $\gamma^0 = \beta$ which was already traceless. From the anti-commutation relation, we know that $\beta\alpha_i = -\alpha_i\beta$ and taking trace on both sides, we get $\mathrm{Tr}\,\gamma^i = 0$

- $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1}$

  Proof. Let us take the spatial indices first. From the definition of α_i and β, we have the following:

$$\gamma^i\gamma^j + \gamma^j\gamma^i = \beta(\alpha_i\beta)\alpha_j + i \leftrightarrow j = -\beta(\beta\alpha_i)\alpha_j + i \leftrightarrow j = -(\alpha_i\alpha_j + \alpha_j\alpha_i) = -2\delta_{ij}\mathbb{1} = 2\eta^{ij}\mathbb{1}$$

[1] More specifically, we should have used $\mathrm{Tr}(A^\dagger B)$

Now, consider one temporal index, that is,

$$\{\gamma^0, \gamma^i\} = \beta^2\alpha_i + \beta\alpha_i\beta = \alpha_i - \alpha_i\beta^2 = 0$$

Finally, we have $\{\gamma^0, \gamma^0\} = 2\beta^2 = 2\mathbb{1} = 2\eta^{00}\mathbb{1}$. Combining the above, we have the identity.

If we instead take the anti-commutation of the covariant indices, then also the identity remains similar, that is,

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}\mathbb{1}$$

- $\mathrm{Tr}(\gamma^\mu\gamma^\nu) = 4\eta^{\mu\nu}$

  Proof. From the anti-commutation relation of the Dirac matrices, we have:

$$2\,\mathrm{Tr}(\gamma^\mu\gamma^\nu) = \mathrm{Tr}\{\gamma^\mu, \gamma^\nu\} = \mathrm{Tr}(2\eta^{\mu\nu}\mathbb{1}) = 2\eta^{\mu\nu} \times 4 \implies \mathrm{Tr}(\gamma^\mu\gamma^\nu) = 4\eta^{\mu\nu}$$

- $(\gamma^5)^2 = \mathbb{1}$

  Proof. From the definition of γ^5 and the anti-commutation between the gamma matrices, we have,

$$\gamma^5\gamma^5 = -(\gamma^0\gamma^1\gamma^2\gamma^3)(\gamma^0\gamma^1\gamma^2\gamma^3) = \gamma^1\gamma^2\gamma^3(\gamma^0)^2\gamma^1\gamma^2\gamma^3 = \gamma^1\gamma^2\gamma^3\gamma^1\gamma^2\gamma^3 = \gamma^2\gamma^3(\gamma^1)^2\gamma^2\gamma^3$$
$$\implies \gamma^5\gamma^5 = -\gamma^2\gamma^3\gamma^2\gamma^3 = \gamma^3(\gamma^2)^2\gamma^3 = -(\gamma^3)^2 = \mathbb{1}$$

- γ^5 is Hermitian.

  Proof. We will use the fact that $\left(\gamma^i\right)^\dagger = -\gamma^i$ for $i = 1, 2, 3$ and $\left(\gamma^0\right)^\dagger = \gamma^0$ along with the anti-commutation relation.

$$(\gamma^5)^\dagger = -i\left(\gamma^3\right)^\dagger\left(\gamma^2\right)^\dagger\left(\gamma^1\right)^\dagger\left(\gamma^0\right)^\dagger = i\gamma^3\gamma^2\gamma^1\gamma^0$$

  To bring γ^0 to correct position (in front), we need 3 flips, for γ^1 we need 2 flips and for γ^2 we need only one flip, hence total $3 + 2 + 1 = 6$ flips. Each flip introduces a negative sign, so total $(-1)^6 = 1$, hence $(\gamma^5)^\dagger = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5$.

- $\{\gamma^5, \gamma^\mu\} = 0$

  Proof. Note that, in $\gamma^5\gamma^\mu = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^\mu$, for exchanging each $\mu \neq \nu$, we introduce a -1. Now, since there are 4 indices, μ must be one of these four indices and hence, we have the same index together, we do not need the -1. Thus, we introduce exactly $(-1)^3 = -1$ while transporting γ^μ to the front. Then we have $\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$, thus proving the identity.

- $\mathrm{Tr}(\gamma^{\mu_1}\gamma^{\mu_2}\ldots\gamma^{\mu_{2n+1}}) = 0$

  Proof. As $(\gamma^5)^2 = \mathbb{1}$, let us multiply γ^5 to the above expression,

$$\gamma^{\mu_1}\gamma^{\mu_2}\ldots\gamma^{\mu_{2n+1}} = \gamma^5\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\ldots\gamma^{\mu_{2n+1}}$$

  Shifting one of the γ^5 to the far right across the $2n + 1$ indices will introduce $(-1)^{2n+1}$ and thus, we get:

$$\gamma^{\mu_1}\gamma^{\mu_2}\ldots\gamma^{\mu_{2n+1}} = (-1)^{2n+1}\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\ldots\gamma^{\mu_{2n+1}}\gamma^5$$

  Taking trace both sides and taking the cyclic property of the trace, we obtain the following. Thus, the product of odd number of gamma matrices is zero.

- $\mathrm{Tr}\left(\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu\right) = 4\eta^{\kappa\lambda}\eta^{\mu\nu} - 4\eta^{\kappa\mu}\eta^{\lambda\nu} + 4\eta^{\kappa\nu}\eta^{\lambda\mu}$

Proof. Using the anti-commutation relation, we note the following:

$$\begin{aligned}\{\gamma^\kappa, \gamma^\lambda\gamma^\mu\gamma^\nu\} &= \gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu + \gamma^\lambda\gamma^\mu\gamma^\nu\gamma^\kappa \\ &= \gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu + \gamma^\lambda\gamma^\kappa\gamma^\mu\gamma^\nu - \gamma^\lambda\gamma^\kappa\gamma^\mu\gamma^\nu + \gamma^\lambda\gamma^\mu\gamma^\nu\gamma^\kappa \\ &= \{\gamma^\kappa, \gamma^\lambda\}\gamma^\mu\gamma^\nu - \gamma^\lambda\gamma^\kappa\gamma^\mu\gamma^\nu - \gamma^\lambda\gamma^\mu\gamma^\kappa\gamma^\nu + \gamma^\lambda\gamma^\mu\gamma^\kappa\gamma^\nu + \gamma^\lambda\gamma^\mu\gamma^\nu\gamma^\kappa \\ &= \{\gamma^\kappa, \gamma^\lambda\}\gamma^\mu\gamma^\nu - \gamma^\lambda\{\gamma^\kappa, \gamma^\mu\}\gamma^\nu + \gamma^\lambda\gamma^\mu\{\gamma^\kappa, \gamma^\nu\} \\ &= 2\eta^{\kappa\lambda}\gamma^\mu\gamma^\nu - 2\eta^{\kappa\mu}\gamma^\lambda\gamma^\nu + 2\eta^{\kappa\nu}\gamma^\lambda\gamma^\mu\end{aligned}$$

Now, note that the required quantity is basically half the trace of the above anti-commutator.

$$\begin{aligned}\operatorname{Tr}\left(\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu\right) = \frac{1}{2}\operatorname{Tr}\{\gamma^\kappa, \gamma^\lambda\gamma^\mu\gamma^\nu\} &= \operatorname{Tr}\left(\eta^{\kappa\lambda}\gamma^\mu\gamma^\nu - \eta^{\kappa\mu}\gamma^\lambda\gamma^\nu + \eta^{\kappa\nu}\gamma^\lambda\gamma^\mu\right) \\ &= \eta^{\kappa\lambda}\operatorname{Tr}\gamma^\mu\gamma^\nu - \eta^{\kappa\mu}\operatorname{Tr}\gamma^\lambda\gamma^\nu + \eta^{\kappa\nu}\operatorname{Tr}\gamma^\lambda\gamma^\mu \\ &= 4\eta^{\kappa\lambda}\eta^{\mu\nu} - 4\eta^{\kappa\mu}\eta^{\lambda\nu} + 4\eta^{\kappa\nu}\eta^{\lambda\mu}\end{aligned}$$

- $\operatorname{Tr}\gamma^5 = 0$

  Proof. Since the Minkowski metric is entire diagonal, $\eta^{ij} = 0$ for $i \neq j$ from which the identity follows using the previous identity.

- $\operatorname{Tr}\left(\gamma^5\gamma^\mu\right) = 0$

  Proof. Note that $\mu \in \{0, 1, 2, 3\}$ and γ^5 contains of all these indices. So, we can bring γ^μ adjacent to the same index in γ^5 by some flips and then it will give $(\gamma^\mu)^2$ which gives ± 1 atmost and the remaining will be a product of odd number of gamma matrices whose trace we had found out to be zero.

- $\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\right) = 0$

  Proof. We will use the fact the γ^5 anti-commutes with every γ^μ. Let us introduce two factors of γ^α in the trace such that $\alpha \neq \mu, \nu$. Note that $(\gamma^\alpha)^2$ will be 1 if $\alpha = 0$, else, it will be -1. In any case, the trace will change by ± 1

$$\begin{aligned}\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\right) = (-1)^{1-\delta_{\alpha,0}}\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\alpha\right) &= (-1)^{1-\delta_{\alpha,0}}(-1)^3\operatorname{Tr}\left(\gamma^\alpha\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\right) \\ &= -(-1)^{1-\delta_{\alpha,0}}\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\alpha\right) \\ &= -\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\right) \\ \implies \operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\right) &= 0\end{aligned}$$

  In the last step, we had used the cyclic property of the trace to bring γ^α to the back again.

- $\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta = -4i\epsilon^{\mu\nu\alpha\beta}$

  Proof. If any two of μ, ν, α, β are same, we could bring them together by flipping, introducing atmost ± 1. The remaining will be the product of γ^5 and two other gamma matrices, whose trace we had already shown to be zero.

  Thus, let us assume that all the indices are different. Then in the product $\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta$, each gamma will appear twice. Hence we flip them such that all same index gamma appear together which would give us one of $\pm\mathbb{1}$, hence the trace will be of the form $\pm i\operatorname{Tr}\mathbb{1} = \pm 4i$.

  Suppose μ, ν, α, β appear as an even permutation of $0, 1, 2, 3$. Then we have,

$$\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right) = -i\operatorname{Tr}\left(\gamma^5 \times i\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right)$$

  Now the red factor can be brought to the form γ^5 using an even number of exchanges. Since we saw that $(\gamma^5)^2 = \mathbb{1}$, we have $\operatorname{Tr}\left(\gamma^5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right) = -4i$.

If μ, ν, α, β appear as an odd permutation of $0, 1, 2, 3$, then the same logic follows, however, the red factor introduces a negative sign since in this case, by an odd number of exchanges, the factor can be brought to the form of γ^5. Hence the trace will give $4i$.

As a summary, we saw that if any two of the four indices are equal, the trace is zero. If indices are even permutation of $0, 1, 2, 3$ then trace $\to -4i$ else if odd permutation, then trace $\to 4i$. This is essentially the same statement as the contravariant Levi-Civita tensor defined as,

$$\epsilon^{\alpha\beta\kappa\lambda} = \begin{cases} +1 & \alpha\beta\kappa\lambda \to \text{ even-permutation of } 0,1,2,3 \\ -1 & \alpha\beta\kappa\lambda \to \text{ odd-permutation of } 0,1,2,3 \\ 0 & \text{else} \end{cases}$$

Thus, we prove the identity.

19.2. Contraction Identities

Let us now look at some contraction identities where the Dirac matrices will be contracted. Unlike the trace identitities which gave a number as a final answer, these will give us some matrices.

- $\gamma^\mu \gamma_\mu = 4\mathbb{1}$

  Proof. Since Dirac matrices can be treated as four vectors, we can lower and raise their indices with the metric η. Then,

$$\begin{aligned} \gamma^\mu \gamma_\mu &= \eta_{\mu\nu}\gamma^\nu\gamma^\mu \\ &= \frac{1}{2}(\eta_{\mu\nu} + \eta_{\nu\mu})\gamma^\nu\gamma^\mu \qquad \text{(as metric is symmetric)} \\ &= \frac{1}{2}(\eta_{\mu\nu}\gamma^\nu\gamma^\mu + \eta_{\nu\mu}\gamma^\nu\gamma^\mu) \\ &= \frac{1}{2}(\eta_{\mu\nu}\gamma^\nu\gamma^\mu + \eta_{\mu\nu}\gamma^\mu\gamma^\nu) \qquad \text{(relabelling dummy indices)} \\ &= \frac{1}{2}\eta_{\mu\nu}(\gamma^\nu\gamma^\mu + \gamma^\mu\gamma^\nu) \\ &= \frac{1}{2}\eta_{\mu\nu} \times 2\eta^{\mu\nu}\mathbb{1} \qquad \text{(from anti-commutation of gamma matrices)} \\ &= \delta^\mu{}_\mu \mathbb{1} \\ &= 4\mathbb{1} \end{aligned}$$

- $\gamma^\mu \gamma^\nu \gamma_\mu = -2\gamma^\nu$

  Proof. Using the anti-commutation relation and the previous identity, we will derive this.

$$\gamma^\mu\gamma^\nu\gamma_\mu = \gamma^\mu\gamma^\nu\eta_{\mu\beta}\gamma^\beta = \gamma^\mu\eta_{\mu\beta}\left(2\eta^{\nu\beta} - \gamma^\beta\gamma^\nu\right) = 2\gamma^\mu\eta_{\mu\beta}\eta^{\nu\beta} - \gamma^\mu\gamma^\beta\eta_{\mu\beta}\gamma^\nu = 2\gamma^\mu\delta^\nu{}_\mu - \underbrace{\gamma^\mu\gamma_\mu}_{4\mathbb{1}}\gamma^\nu = -2\gamma^\nu$$

- $\gamma^\mu \gamma_\nu \gamma_\lambda \gamma_\mu = 4\eta_{\lambda\nu}\mathbb{1}$

  Proof. We will use the anti-commutation relation and the previous identity together to reduce the expression.

$$\begin{aligned} \gamma^\mu\gamma_\nu\gamma_\lambda\gamma_\mu = \gamma^\mu\gamma_\nu(2\eta_{\lambda\mu} - \gamma_\mu\gamma_\lambda) = 2\eta_{\lambda\mu}\gamma^\mu\gamma_\nu - \gamma^\mu\gamma_\nu\gamma_\mu\gamma_\lambda = 2\gamma_\lambda\gamma_\nu - \eta_{\alpha\nu}\underbrace{\gamma^\mu\gamma^\alpha\gamma_\mu}_{-2\gamma^\alpha}\gamma_\lambda &= 2(\gamma_\lambda\gamma_\nu + \gamma_\nu\gamma_\lambda) \\ &= 4\eta_{\lambda\nu}\mathbb{1} \end{aligned}$$

- $\gamma^\mu \gamma_\nu \gamma_\lambda \gamma_\rho \gamma_\mu = -2\gamma_\rho\gamma_\lambda\gamma_\nu$

Proof. We will again use the anti-commutation relation and the previous identity for this.

$$\gamma^\mu\gamma_\nu\gamma_\lambda\gamma_\rho\gamma_\mu = \gamma^\mu\gamma_\nu\gamma_\lambda(2\eta_{\rho\mu} - \gamma_\mu\gamma_\rho) = 2\gamma_\rho\gamma_\nu\gamma_\lambda - \underbrace{\gamma^\mu\gamma_\nu\gamma_\lambda\gamma_\mu}_{4\eta_{\lambda\nu}\mathbb{1}}\gamma_\rho = 2\gamma_\rho(\gamma_\nu\gamma_\lambda - 2\eta_{\lambda\nu}) = -2\gamma_\rho \times (\underbrace{2\eta_{\lambda\nu} - \gamma_\nu\gamma_\lambda}_{\gamma_{\lambda\nu}})$$
$$= -2\gamma_\rho\gamma_\lambda\gamma_\nu$$

These are few of the contraction and trace identities that we checked. There exists many such other identities and we can just play with the Dirac matrices to find them 😉!

19.3. Orthonormal Basis

In the Lec 18, we had stated a basis set for the set of 4×4 matrices, however, we have not explicitly shown that it is a basis. Since we are equipped with the traces (which form the inner product for the space) we are now in a position to show that. To show that the set is an orthonormal basis, we need to show two things:

- Linear independence of the basis elements.
- orthognality of the basis elements with respect to the inner product.

Let us start by showing the orthognality. For that, we should painstakingly calculate the traces between the elements in the basis set. However, calculation for only $\binom{5}{2} + 5 = 15$ traces need to be shown as trace is symmetric.

- Term $\longrightarrow \mathrm{Tr}(\mathbb{1}\cdot\Gamma^\alpha)$ where $\Gamma^\alpha \in \Gamma$

  We have $\mathrm{Tr}(\mathbb{1}\cdot\mathbb{1}) = 4$ obviously. From the above derived trace relations, we know $\mathrm{Tr}\big(\mathbb{1}\cdot\gamma^5\big) = 0, \mathrm{Tr}\big(\mathbb{1}\cdot\gamma^\mu\gamma^5\big) = 0, \mathrm{Tr}(\mathbb{1}\cdot\gamma^\mu) = 0$ and

  $$\mathrm{Tr}(\mathbb{1}\cdot\Sigma^\mu\gamma^\nu) = \frac{i}{2}(4\eta^{\mu\nu} - 4\eta^{\nu\mu}) = 0$$

- Term $\longrightarrow \mathrm{Tr}(\gamma^\mu\cdot\Gamma^\alpha)$ where $\Gamma^\alpha \in \Gamma$

  From the trace identity, we have $\mathrm{Tr}(\gamma^\mu\gamma^\nu) = 4\eta^{\mu\nu}$ and $\mathrm{Tr}(\gamma^\mu) = 0$. $\mathrm{Tr}\big(\gamma^\mu\gamma^5\big) = 0$ as this is a product of five gamma matrices. $\mathrm{Tr}(\gamma^\mu\Sigma^{\nu\alpha}) = 0$ since it contains the product of three gamma matrices. $\mathrm{Tr}\big(\gamma^\mu\gamma^\nu\gamma^5\big) = 0$ from the trace identity.

- Term $\longrightarrow \mathrm{Tr}\big(\gamma^5\cdot\Gamma^\alpha\big)$ where $\Gamma^\alpha \in \Gamma$

  We have $\mathrm{Tr}\big(\gamma^5\gamma^\mu\gamma^5\big) = -\mathrm{Tr}\big((\gamma^5)^2\gamma^\mu\big) = 0, \mathrm{Tr}\big(\gamma^5\Sigma^{\mu\nu}\big) = 0$ as $\mathrm{Tr}\big(\gamma^\mu\gamma^\nu\gamma^5\big) = 0$ and $\mathrm{Tr}\big((\gamma^5)^2\big) = 4$

- Term $\longrightarrow \mathrm{Tr}\big(\gamma^\mu\gamma^5\cdot\Gamma^\alpha\big)$ where $\Gamma^\alpha \in \Gamma$

  We have $\mathrm{Tr}\big(\gamma^\mu\gamma^5\gamma^\nu\gamma^5\big) = -\mathrm{Tr}(\gamma^\mu\gamma^\nu) = -4\eta^{\mu\nu}$ and $\mathrm{Tr}\big(\gamma^\mu\gamma^5\Sigma^{\alpha\rho}\big) = 0$ as product of seven gamma matrices.

- Term $\longrightarrow \mathrm{Tr}\big(\Sigma^{\alpha\beta}\cdot\Gamma^\alpha\big)$ where $\Gamma^\alpha \in \Gamma$

  As an exercise, let us calculate this trace explicitly from the definition itself.

  $$\mathrm{Tr}\,\Sigma^{\alpha\beta}\Sigma^{\mu\nu} = -\frac{1}{4}\mathrm{Tr}\Big((\gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha)(\gamma^\mu\gamma^\nu - \gamma^\beta\gamma^\mu)\Big) = -\frac{1}{4}\mathrm{Tr}\Big(\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu\Big) + \frac{1}{4}\mathrm{Tr}\Big(\gamma^\alpha\gamma^\beta\gamma^\nu\gamma^\mu\Big)$$
  $$+\frac{1}{4}\mathrm{Tr}\Big(\gamma^\beta\gamma^\alpha\gamma^\mu\gamma^\nu\Big) - \frac{1}{4}\mathrm{Tr}\Big(\gamma^\beta\gamma^\alpha\gamma^\nu\gamma^\mu\Big)$$

  Using the trace identities, this elongates to:

  $$-(\eta^{\alpha\beta}\eta^{\mu\nu} - \eta^{\alpha\mu}\eta^{\beta\nu} + \eta^{\alpha\nu}\eta^{\beta\mu}) + (\eta^{\alpha\beta}\eta^{\nu\mu} - \eta^{\alpha\nu}\eta^{\beta\mu} + \eta^{\alpha\mu}\eta^{\beta\nu})$$
  $$+ (\eta^{\beta\alpha}\eta^{\mu\nu} - \eta^{\beta\mu}\eta^{\alpha\nu} + \eta^{\beta\nu}\eta^{\alpha\mu}) - (\eta^{\beta\alpha}\eta^{\nu\mu} - \eta^{\beta\nu}\eta^{\alpha\mu} + \eta^{\beta\mu}\eta^{\alpha\nu})$$

  Note that there are four of each blue and green colored terms while the red terms cancel and hence the trace becomes:

  $$\mathrm{Tr}\,\Sigma^{\alpha\beta}\Sigma^{\mu\nu} = 4(\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu})$$

We thus see that for $\Gamma^a, \Gamma^b \in \Gamma$, $\text{Tr}\left(\Gamma^a\Gamma^b\right) \neq 0$ only when $\Gamma^a = \Gamma^b$ where the equality means that both are same kind of matrices in the Clifford algebra basis. We want to find some kind of definite normalisation which does not depend on the sign of the trace. For that, we will change the basis element $\gamma^\mu\gamma^5 \to i\gamma^\mu\gamma^5$ and we will lower the index, Γ_a, by lowering all the indices of Γ^a using $\eta_{\mu\nu}$. Then, we will find that:

$$\text{Tr}\left(\Gamma_a\Gamma^b\right) = 4\delta^b{}_a$$

Using this definition, the linear independence can be easily shown. Suppose we have:

$$\sum_a C_a\Gamma^a = 0$$

Multiplying the equation both sides with Γ^b and then taking the trace, only the term with C_b will survive on the LHS and will be zero. Thus, $C_a = 0 \ \forall \ a$ which implies that the elements are linearly independent.

Since this forms a basis, we can expand any 4×4 matrix in terms of these elements. Thus,

$$M = \sum_a C_a^M \Gamma^a$$

Then multiplying by another element and taking trace, we get:

$$\text{Tr}\,\Gamma_b M = \sum_a C_a^M \,\text{Tr}\,\Gamma_b\Gamma^a = \sum_a 4C_a^M\delta^a{}_b = 4C_b^M \implies \boxed{C_b^M = \frac{1}{4}\,\text{Tr}\,\Gamma_b M}$$

And hence the matrix is expanded as:

$$M = \sum_a \frac{1}{4}\,\text{Tr}(\Gamma_a M)\ \Gamma^a \tag{12}$$

Now let us explicitly introduce the matrix indices which we have suppressed till now. Note that there are two kind of indices here, one is the a index which tells us what type of the basis element it is and now we introduce the matrix indices, which will run from $1, 2, 3, 4$. So, Γ^a_{12} actually means the $(1, 2)$ index of the matrix of type Γ^a. Using matrix indices, the above equation becomes:

$$M_{\alpha\beta} = \frac{1}{4}\sum_a (\Gamma_a)_{\eta\lambda} M_{\lambda\eta}\Gamma^a{}_{\alpha\beta} = \frac{1}{4}\sum_a \left\{(\Gamma_a)_{\eta\lambda}(\Gamma^a)_{\alpha\beta}\right\} M_{\lambda\eta} \implies \boxed{\frac{1}{4}\sum_a (\Gamma_a)_{\eta\lambda}(\Gamma^a)_{\alpha\beta} = \delta_{\alpha\lambda}\delta_{\beta\eta}} \tag{13}$$

This defines a *completeness relation* for these matrices and will be later used to derive some important identitites.

Lecture 20: Transformation of Spinors

To understand how spinors actually transform under Lorentz transformation, we first need to have an idea about what spinors are. Spinors are basically another kind of representation of the Lorentz group loosely speaking. There are different concepts of group and representation theory involved which lead to many different subtleties, however, we shall not digress much into those subtleties and obtain only those ideas which are absolutely needed.

20.1. Spinor Representation of Lorentz group

We know that any representation of the Lorentz group, say $\text{S}[\Lambda]$ can be written in terms of the rotation generators $\mathbf{J}$ and the boost generators $\mathbf{K}$ (from A):

$$\text{S}[\Lambda] = \exp\left(-i\ \boldsymbol{\theta}\cdot\mathbf{J} + i\boldsymbol{\eta}\cdot\mathbf{K}\right)$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ are the rotation and boost parameters. The normal boost and rotation generators are not closed under commutation, however if we define[1]:

$$\mathbf{J}^\pm := \frac{\mathbf{J} \pm i\mathbf{K}}{2}$$

[1] This is called *complexification*

Then, one can check that the Lorentz algebra becomes:

$$\begin{aligned} \left[J^{+,i}, J^{+,j}\right] &= i\varepsilon^{ijk} J^{+,k} \\ \left[J^{-,i}, J^{-,j}\right] &= \varepsilon^{ijk} J^{-,k} \\ \left[J^{+,i}, J^{-,j}\right] &= 0 \end{aligned}$$

Now, this is familiar to the $SU(2)$ algebra that we had seen from, say the *angular momentum*. In fact, from the complexification of the generators we see two copies of the angular momentum algebra (so we can write $\mathfrak{so}(1,3) = \mathfrak{su}(2) \oplus \mathfrak{su}(2)$) and we can thus label the representations by (j_-, j_+) where $j_\pm \in \{0, \frac{1}{2}, 1, \frac{3}{2} \ldots\}$. We will briefly focus on the most popular representations, namely:

- *Scalar representation* $\longrightarrow (0,0)$
- *Left-handed Weyl spinor representation* $\longrightarrow (1/2, 0)$
- *Right-handed Weyl spinor representation* $\longrightarrow (0, 1/2)$
- *Dirac spinor representation* $\longrightarrow (0, 1/2) \oplus (1/2, 0)$

The scalar representation is the trivial one-dimensional representation where $\mathbf{J}^\pm = 0 \implies \mathbf{J} = 0, \mathbf{K} = 0$ and hence is of not much interest to us. All but the Dirac representation are irreducible representations. Let us now focus on the Weyl representation.

Weyl Representation

Let us consider the left handed spinor representation. The left-handed spinor itself can be denoted by ψ_{L} while any representation of the Lorentz group can be represented by Λ_{L}. In this case, $\mathbf{J}^+ = 0$ while $\mathbf{J}^-$ resembles the representation of spin-half particles and can be taken to be $\mathbf{J}^- \to \frac{\boldsymbol{\sigma}}{2}$ where $\boldsymbol{\sigma}$ is the Pauli vector. Then from the definition,

$$\begin{aligned} \mathbf{J} &= \mathbf{J}^+ + \mathbf{J}^- = 0 + \frac{\boldsymbol{\sigma}}{2} = \frac{\boldsymbol{\sigma}}{2} \\ \mathbf{K} &= -i(\mathbf{J}^+ - \mathbf{J}^-) = i\frac{\boldsymbol{\sigma}}{2} \end{aligned}$$

Note that, the Pauli vector is Hermitian and then so is $\mathbf{J}$ but $\mathbf{K}$ is anti-Hermitian instead, so overall the Weyl representation is not a unitary representation. This is explained by the fact that the Lorentz group is not a *compact* group and by a theorem, any non-trivial finite dimensional irreducible representation cannot be unitary.

Thus, the representation of the Lorentz group and the transformation of a left-handed Weyl spinor becomes:

$$\mathrm{S}[\Lambda_{\mathrm{L}}] = \exp\left((-i\boldsymbol{\theta} - \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) \qquad \psi_{\mathrm{L}} \to \Lambda_{\mathrm{L}} \psi_{\mathrm{L}}$$

Using a similar argument, we can see that the right-handed Weyl spinor is characterised by:

$$\mathrm{S}[\Lambda_{\mathrm{R}}] = \exp\left((-i\boldsymbol{\theta} + \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) \qquad \psi_{\mathrm{R}} \to \Lambda_{\mathrm{R}} \psi_{\mathrm{R}}$$

Dirac representation

Now, let us come to the Dirac representation which is a reducible representation and can be written as[1]:

$$\Lambda_{\mathrm{D}} \equiv \begin{pmatrix} \Lambda_{\mathrm{L}} & \\ & \Lambda_{\mathrm{R}} \end{pmatrix} \qquad \Psi \to \Lambda_{\mathrm{D}} \Psi$$

▷ Why is the Dirac representation even necessary when we already had the Weyl spinors?

Well, for that consider the parity transformation where $\mathbf{x} \to -\mathbf{x}$. In this case, $\mathbf{K} \to -\mathbf{K}$ since the

[1] We will denote Dirac spinors with capital letters, that is, Ψ while Weyl spinors with small letters, ψ to distinguish between them, since these are inequivalent representations. Of course, it might not always be possible to remember this and the notation might slip sometimes 😢

boost velocity is reversed, however, $\mathbf{J}$ remains same (and thus acts as a *pseudovector*) as intuitively, both position and velocity is reversed.

Thus, under parity $\mathbf{J}^{\pm} \to \mathbf{J}^{\mp}$ and an object in the (j_-, j_+) representation is changed to an object in the (j_+, j_-) representation which is different unless $j_+ = j_-$.

Then, the Weyl spinors (objects of kind $(0, 1/2)$ and $(1/2, 0)$) individually cannot be the basis for a representation of the parity transformation. In this case, it is convenient to work with fields which provide a representation of Lorentz and parity transformations together. Thus, we consider a Dirac spinor written in terms of two Weyl spinors[1],

$$\Psi = \begin{pmatrix} \psi_{\mathrm{L}} \\ \psi_{\mathrm{R}} \end{pmatrix}$$

which is a $(0, 1/2) \oplus (1/2, 0)$ object and conforms to parity transformations. Consider an infinitesimal transformation of the Dirac representation, which can be equivalently written as the infinitesimal transformation of the two Weyl representations.

$$\exp\left((-i\boldsymbol{\theta} - \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) \approx \mathbb{1} + \left((-i\boldsymbol{\theta} - \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) = \mathbb{1} - \frac{i}{2}\theta^i\sigma^i - \frac{1}{2}\eta^i\sigma^i$$
$$\exp\left((-i\boldsymbol{\theta} + \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) \approx \mathbb{1} + \left((-i\boldsymbol{\theta} + \boldsymbol{\eta}) \cdot \frac{\boldsymbol{\sigma}}{2}\right) = \mathbb{1} - \frac{i}{2}\theta^i\sigma^i + \frac{1}{2}\eta^i\sigma^i$$

Substituting this in the Dirac representation, we obtain:

$$\begin{aligned}
\Lambda_{\mathrm{D}} &= \mathbb{1} - \frac{i}{2}\theta^i \begin{pmatrix} \sigma^i & \\ & \sigma^i \end{pmatrix} - \frac{1}{2}\eta^i \begin{pmatrix} \sigma^i & \\ & -\sigma^i \end{pmatrix} \\
&= \mathbb{1} - \frac{i}{4}\varepsilon^{ijk}\Omega_{jk} \begin{pmatrix} \sigma^i & \\ & \sigma^i \end{pmatrix} - \frac{1}{2}\Omega^{i0} \begin{pmatrix} \sigma^i & \\ & -\sigma^i \end{pmatrix} \\
&= \mathbb{1} - \frac{i}{4}\Omega_{jk}\Sigma^{jk} - \frac{i}{2}\Omega^{i0}\Sigma^{0i} \\
&= \mathbb{1} - \frac{i}{4}\Omega_{jk}\Sigma^{jk} + \frac{i}{2}\Omega_{i0}\Sigma^{0i} \\
&= \mathbb{1} - \frac{i}{4}\Omega_{jk}\Sigma^{jk} - \frac{i}{2}\Omega_{0i}\Sigma^{0i} \\
&= \mathbb{1} - \frac{i}{4}\Omega_{jk}\Sigma^{jk} - \frac{i}{4}\Omega_{0i}\Sigma^{0i} - \frac{i}{4}\Omega_{i0}\Sigma^{i0} \\
&= \mathbb{1} - \frac{i}{4}\Omega_{\rho\sigma}\Sigma^{\rho\sigma}
\end{aligned}$$

where we have used the form of the matrix $\Sigma^{\mu\nu}$ in the chiral basis[2] (since we are working with Weyl representation where the representations Λ act on the Weyl spinors, the chiral basis is an appropriate choice). We had previously found the form of Σ^{jk} in 18.2 which will not change in the chiral basis. We now just find the form of Σ^{0i}:

$$\begin{aligned}
-2i\Sigma^{0i} &= \beta\gamma^i - \gamma^i\beta \\
&= \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \\
&= \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} - \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \\
&= -2\begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}
\end{aligned}$$

[1] A subtlety is that this is only true for the *chiral basis* where γ^0 takes an off-diagonal form

[2] Only γ^0 has a different form in the chiral basis.

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

Thus, we found an expression for the infinitesimal transformation of the spinor generators. Hence, any Dirac spinor representation of the Lorentz group can be found by exponentiating these generators:

$$\mathrm{S}[\Lambda] = \exp\left(-\frac{i}{4}\Omega_{\rho\sigma}\Sigma^{\rho\sigma}\right)$$

As specified earlier, S[Λ] cannot be unitary in general, since Lorentz group is not a compact group and as a result of which, the obvious choice for the inner product of a Dirac spinor is not *Lorentz invariance*. We can explicitly see this as,

$$\Psi'^{\dagger}\Psi' = \Psi^{\dagger} \underbrace{\mathrm{S}[\Lambda]^{\dagger}\mathrm{S}[\Lambda]}_{\neq \mathbb{1}} \Psi \neq \Psi^{\dagger}\Psi$$

20.2. Constructing Bilinears

We saw that the obvious notion of an *adjoint* is not Lorentz invariant which is problematic. We will now try to construct some tensorial quantitites out of the spinors in a different way, which will be Lorentz invariant/covariant. For that, we need to look at some properties.

Firstly note that γ^i's are anti-Hermitian while γ^0 is Hermitian. Then we have:

$$\gamma^0\gamma^i\gamma^0 = (-\gamma^0)^2\gamma^i = -\gamma^i = \gamma^{i\dagger} \qquad \gamma^0\gamma^0\gamma^0 = \gamma^0 = \gamma^{0\dagger}$$

Hence we have the identity:

$$\boxed{\gamma^0\gamma^\mu\gamma^0 = \gamma^{\mu\dagger}} \tag{14}$$

Now, we will try to see how the generators transform under the adjoint. For that, note:

$$\begin{aligned}
\Sigma^{\mu\nu\dagger} &= \left(\frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right]\right)^\dagger \\
&= -\frac{i}{2}\left[\gamma^{\nu\dagger}, \gamma^{\mu\dagger}\right] \\
&= -\frac{i}{2}\{(\gamma^0\gamma^\nu\gamma^0)(\gamma^0\gamma^\mu\gamma^0) - (\gamma^0\gamma^\mu\gamma^0)(\gamma^0\gamma^\nu\gamma^0)\} \\
&= -\frac{i}{2}\gamma^0\{\gamma^\nu\gamma^\mu - \gamma^\mu\gamma^\nu\}\gamma^0 \\
&= -\frac{i}{2}\gamma^0\left[\gamma^\nu, \gamma^\mu\right]\gamma^0 \\
&= -\gamma^0\Sigma^{\nu\mu}\gamma^0 \\
&= \gamma^0\Sigma^{\mu\nu}\gamma^0
\end{aligned}$$

Now note the property of the generators:

$$\gamma^0 \underbrace{\left(-\frac{i}{4}\Omega_{\rho\sigma}\Sigma^{\rho\sigma}\right)}_{\text{generator of } \mathrm{S}[\Lambda]} \gamma^0 = \frac{i}{4}\Omega_{\rho\sigma}\gamma^0\Sigma^{\rho\sigma\dagger}\gamma^0 = + \underbrace{\frac{i}{4}\Omega_{\rho\sigma}\Sigma^{\rho\sigma}}_{\text{generator of } \mathrm{S}^{-1}[\Lambda]} \qquad \text{from the previous property}$$

Translating this to the actual Lorentz representations, we obtain:

$$\gamma^0\ \mathrm{S}^\dagger[\Lambda]\ \gamma^0 = \mathrm{S}^{-1}[\Lambda]\ \longleftrightarrow \mathrm{S}^\dagger[\Lambda] = \gamma^0\ \mathrm{S}^{-1}[\Lambda]\ \gamma^0$$

From this, we can actually define a new kind of adjoint:

$$\overline{\Psi} := \Psi^\dagger\gamma^0$$

Let us now check how this quantity transforms under a Lorentz transformation

$$\overline{\Psi}(x) \longrightarrow \overline{\Psi}'(x') = (\mathrm{S}[\Lambda]\Psi(x))^\dagger\gamma^0 = \Psi^\dagger(x)\mathrm{S}^\dagger[\Lambda]\gamma^0 = \Psi^\dagger(x)\gamma^0\mathrm{S}^{-1}[\Lambda] = \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]$$

Hence this *Dirac adjoint* transforms with a copy of $\mathrm{S}^{-1}[\Lambda]$. Using these transformations, we can construction a few tensorial quantitites. For that, remember the basis Γ that we had constructed for 4×4 matrices. Let us construct what we call *bilinears*, as $\overline{\Psi}\ (\ \cdot\ )\ \Psi$. These are called so since these are linear in both Ψ and $\overline{\Psi}$.

- **Basis element**: $\mathbb{1}$

$$\overline{\Psi}\ \mathbb{1}\ \Psi = \overline{\Psi}\Psi \longrightarrow \overline{\Psi}'\Psi' = \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\ \mathrm{S}[\Lambda]\Psi = \overline{\Psi}\Psi$$

  Thus, this quantity is a Lorentz *scalar* since it is invariant under transformation. We denote it by $\mathcal{S}$.

- **Basis element**: γ^μ

$$\begin{aligned}\overline{\Psi}\ \gamma^\mu\ \Psi \longrightarrow \overline{\Psi}'\gamma^\mu\Psi' &= \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\gamma^\mu \mathrm{S}[\Lambda]\ \Psi \\ &= \overline{\Psi}\ \Lambda^\mu{}_\nu\gamma^\nu\ \Psi \quad \text{(using Eq. 10)} \\ &= \Lambda^\mu{}_\nu\big(\overline{\Psi}\ \gamma^\nu\ \Psi\big)\end{aligned}$$

  This quantity transform with one copy of $\Lambda^\mu{}_\nu$ and hence is a *contravariant vector*. We denote it by $\mathcal{V}$

- **Basis element**: $\Sigma^{\mu\nu}$

$$\begin{aligned}\overline{\Psi}\ \Sigma^{\mu\nu}\ \Psi \longrightarrow \overline{\Psi}'\ \Sigma^{\mu\nu}\ \Psi' &= \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\Sigma^{\mu\nu}\mathrm{S}[\Lambda]\ \Psi \\ &= \frac{i}{2}\overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\,[\gamma^\mu,\gamma^\nu]\,\mathrm{S}[\Lambda]\Psi \\ &= \frac{i}{2}\big\{\overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\gamma^\mu\gamma^\nu\mathrm{S}[\Lambda]\Psi - \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\gamma^\nu\gamma^\mu\mathrm{S}[\Lambda]\Psi\big\} \\ &= \frac{i}{2}\big\{\overline{\Psi}\ \big(\mathrm{S}^{-1}[\Lambda]\gamma^\mu\mathrm{S}[\Lambda]\big)\big(\mathrm{S}^{-1}[\Lambda]\gamma^\nu\mathrm{S}[\Lambda]\big)\Psi - \overline{\Psi}\ \big(\mathrm{S}^{-1}[\Lambda]\gamma^\nu\mathrm{S}[\Lambda]\big)\big(\mathrm{S}^{-1}[\Lambda]\gamma^\mu\mathrm{S}[\Lambda]\big)\Psi\big\} \\ &= \frac{i}{2}\Big\{\overline{\Psi}\Big(\Lambda^\mu{}_\alpha\gamma^\alpha\Lambda^\nu{}_\beta\gamma^\beta - \Lambda^\nu{}_\beta\gamma^\beta\Lambda^\mu{}_\alpha\gamma^\alpha\Big)\Psi\Big\} \\ &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\frac{i}{2}\overline{\Psi}\left[\gamma^\alpha,\gamma^\beta\right]\Psi \\ &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\Big(\overline{\Psi}\ \Sigma^{\alpha\beta}\ \Psi\Big)\end{aligned}$$

  This quantity transforms with two copies of Λ and hence, acts as a rank-2 contravariant tensor. We denote it by $\mathcal{T}^{\mu\nu}$

- **Basis element**: γ^5

  γ^5 is defined in the following way using the Levi-Civita tensor:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\varepsilon_{\alpha\beta\lambda\delta}\gamma^\alpha\gamma^\beta\gamma^\lambda\gamma^\delta$$

$$\overline{\Psi}\ \gamma^5\ \Psi =\!\!\longrightarrow \overline{\Psi}'\ \gamma^5\ \Psi' = \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\gamma^5\mathrm{S}[\Lambda]\ \Psi = -\frac{i}{4!}\varepsilon_{\alpha\beta\lambda\delta}\overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\ \gamma^\alpha\gamma^\beta\gamma^\lambda\gamma^\delta\mathrm{S}[\Lambda]\ \Psi$$

  Now, what we did earlier, by introducing $\mathrm{S}[\Lambda]\mathrm{S}^{-1}[\Lambda]$ between each Dirac matrices, we can introduce a copy of Λ. Doing this, we obtain:

$$\overline{\Psi}'\ \gamma^5\ \Psi' = -\frac{i}{4!}\ \overline{\Psi}\varepsilon_{\alpha\beta\lambda\delta}\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu\Lambda^\lambda{}_\rho\Lambda^\delta{}_\theta\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\theta\ \Psi$$

Fun Fact:
The determinant of a square $n \times n$ matrix can be written in terms of the Levi-Civita tensor as:

$$\det(A) = \frac{1}{n!}\varepsilon_{\mu_1\mu_2\ldots\mu_n}\varepsilon^{\nu_1\nu_2\ldots\nu_n}A^{\mu}{}_{1\nu 1}A^{\mu}{}_{2\nu 2}\ldots A^{\mu}{}_{n\nu n}$$

Contracting the above equation by $\varepsilon_{\nu_1\nu_2\ldots\nu_n}$ we obtain:

$$\varepsilon_{\nu_1\nu_2\ldots\nu_n}\det(A) = \varepsilon_{\mu_1\mu_2\ldots\mu_n}A^{\mu_1}{}_{\nu_1}A^{\mu_2}{}_{\nu_2}\ldots A^{\mu_n}{}_{\nu_n}$$

Using the above fact for Λ, we get:

$$\begin{aligned}\overline{\Psi}'\ \gamma^5\ \Psi' &= -\frac{i}{4!}\overline{\Psi}\ \varepsilon_{\mu\nu\rho\theta}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\theta\det(\Lambda)\ \Psi \\ &= \det(\Lambda)(\overline{\Psi}\ \gamma^5\ \Psi)\end{aligned}$$

This quantity changes almost like a scalar but with the determinant of the transformation. For proper Lorentz element, $\det(\Lambda) = 1$ and hence this is invariant, however, a minus sign is introduced for improper Lorentz transformation. Hence this is instead called a *pseudo-scalar* and is denoted by $\mathcal{P}$

- **Basis element**: $\gamma^\mu\gamma^5$

$$\begin{aligned}\overline{\Psi}\ \gamma^\mu\gamma^5\ \Psi \Longrightarrow \overline{\Psi}'\ \gamma^\mu\gamma^5\ \Psi' = \overline{\Psi}\ \mathrm{S}^{-1}[\Lambda]\gamma^\mu\gamma^5\mathrm{S}[\Lambda]\ \Psi &= \overline{\Psi}\ (\mathrm{S}^{-1}[\Lambda]\gamma^\mu\mathrm{S}[\Lambda])(\mathrm{S}^{-1}[\Lambda]\gamma^5\mathrm{S}[\Lambda])\ \Psi \\ &= \overline{\Psi}\ \Lambda^\mu{}_\nu\gamma^\nu \times \det(\Lambda)\gamma^5\ \Psi \\ &= \det(\Lambda)\Lambda^\mu{}_\nu(\overline{\Psi}\ \gamma^\nu\gamma^5\ \Psi)\end{aligned}$$

This quantity transforms very close to a rank-2 tensor but with the determinant of the transformation and hence changes sign under improper Lorentz transformation. Hence it is called a *pseudo-vector* or *axial vector* and is denoted by $\mathcal{A}$

These five bilinears, corresponding to each element of the basis of 4×4 matrices, is thus useful in constructing general quantities transforming properly under the Lorentz transformation.

Lecture 21: Fierz Identities

In the previous section we saw how different *bilinears* can be created through the use of the basis elements of the Clifford algebra. We also saw the transformation of these quantities and accordingly specified names to them viz. scalars, pseudoscalars, vector, axial vector and tensor. In this section, we will find an identity that will allow us to rearrange the bilinears of the product of two spinors in a different ordering. These can be useful to write a given interaction in some physically meaningful form.

21.1. Outer product of Spinors

Let us consider two Dirac spinors Ψ_1 and Ψ_2. We previously considered terms of type $\overline{\Psi}_1\Psi_2 = \overline{\Psi}_1\ \gamma^0\ \Psi_2$ which is essentially the product of a row vector and a column vector, producing a number .

Let us now consider the *outer product* between them which will result in some matrix. Since the Clifford algebra produces a basis of 4×4 matrices, we can write the outer product as a linear combination of the elements of the basis set Γ. We thus have:

$$\Psi_2\overline{\Psi}_1 = a\mathbb{1} + b_\mu\gamma^\mu + c_{\mu\nu}\Sigma^{\mu\nu} + d_\mu\gamma^\mu\gamma^5 + f\gamma^5$$

We just need to find these coefficients. The general technique will be to multiple the above equation with some element $\Gamma^a \in \Gamma$ and then take the trace. Since the basis elements are orthogonal, the trace will give zero for dissimilar elements and only contribution is from the same basis element. Let us start!

- Multiply by $\mathbb{1}$ and take the trace:

$$\operatorname{Tr}\left(\mathbb{1}\Psi_2\overline{\Psi}_1\right) = a\operatorname{Tr}(\mathbb{1}) = 4a \implies a = \frac{1}{4}\operatorname{Tr}\left(\Psi_2\overline{\Psi}_1\right) = \frac{1}{4}\overline{\Psi}_1\Psi_2 \longrightarrow \frac{1}{4}\mathcal{S}_{12}$$

  where $\mathcal{S}_{12}$ is the notation to denote the scalar bilinear with spinors Ψ_1 and Ψ_2.

- Multiply by γ^α and take the trace:

$$\operatorname{Tr}\left(\gamma^\alpha\Psi_2\overline{\Psi}_1\right) = b_\mu\operatorname{Tr}(\gamma^\alpha\gamma^\mu) = 4\eta^{\alpha\mu}b_\mu \implies b^\alpha = \frac{1}{4}\operatorname{Tr}\left(\gamma^\alpha\Psi_2\overline{\Psi}_1\right) = \frac{1}{4}\overline{\Psi}_1\ \gamma^\alpha\ \Psi_2 \longrightarrow \frac{1}{4}\mathcal{V}^\alpha_{12}$$

- Multiply by γ^5 and take the trace:

$$\operatorname{Tr}\left(\gamma^5\Psi_2\overline{\Psi}_1\right) = f\operatorname{Tr}\left((\gamma^5)^2\right) = 4f \implies f = \frac{1}{4}\operatorname{Tr}\left(\gamma^5\Psi_2\overline{\Psi}_1\right) = \frac{1}{4}\overline{\Psi}_1\ \gamma^5\ \Psi_2 \longrightarrow \frac{1}{4}\mathcal{P}_{12}$$

- Multiply by $\gamma^\alpha\gamma^5$ and take the trace:

$$\operatorname{Tr}\left(\gamma^\alpha\gamma^5\Psi_2\overline{\Psi}_1\right) = d_\mu\operatorname{Tr}\left(\gamma^\alpha\gamma^5\gamma^\mu\gamma^5\right) = -d_\mu\operatorname{Tr}\left(\gamma^\alpha\gamma^\mu(\gamma^5)^2\right) = -d_\mu\operatorname{Tr}(\gamma^\alpha\gamma^\mu) = -4\eta^{\alpha\mu}d_\mu$$
$$\implies d^\alpha = -\frac{1}{4}\operatorname{Tr}\left(\gamma^\alpha\gamma^5\Psi_2\overline{\Psi}_1\right) = -\frac{1}{4}\overline{\Psi}_1\ \gamma^\alpha\gamma^5\ \Psi_2 \longrightarrow -\frac{1}{4}\mathcal{A}^\alpha_{12}$$

- Multiply by $\Sigma^{\alpha\beta}$ and take the trace:

$$\operatorname{Tr}\left(\Sigma^{\alpha\beta}\Psi_2\overline{\Psi}_1\right) = c_{\mu\nu}\operatorname{Tr}\left(\Sigma^{\alpha\beta}\Sigma^{\mu\nu}\right)$$

  From Lec 19, we have

$$\operatorname{Tr}\Sigma^{\alpha\beta}\Sigma^{\mu\nu} = 4(\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu})$$

  Substituting this in the actual term, we obtain:

$$\operatorname{Tr}\left(\Sigma^{\alpha\beta}\Psi_2\overline{\Psi}_1\right) = 4c_{\mu\nu}\left(\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}\right) = 4(c^{\alpha\beta} - c^{\beta\alpha}) = 8c^{\alpha\beta} \implies c^{\alpha\beta} = \frac{1}{8}\overline{\Psi}_1\ \Sigma^{\alpha\beta}\ \Psi_2 \longrightarrow \frac{1}{8}\mathcal{T}^{\alpha\beta}_{12}$$

  where we used the fact that $c_{\mu\nu}$ is anti-symmetric since $\Sigma^{\mu\nu}$ are also anti-symmetric.

We found all the components and the outer product of two Dirac spinors is represented by:

$$\boxed{\Psi_2\overline{\Psi}_1 = \frac{1}{4}\left\{\mathcal{S}_{12}\mathbb{1} + (\mathcal{V}_{12})_\mu\gamma^\mu + \frac{1}{2}(\mathcal{T}_{12})_{\mu\nu}\Sigma^{\mu\nu} - (\mathcal{A}_{12})_\mu\gamma^\mu\gamma^5 + \mathcal{P}_{12}\gamma^5\right\}}$$

21.2. Rearranging Spinors

Let $\Psi_1, \Psi_2, \Psi_3, \Psi_4$ be four Dirac spinors given to us and consider $(\overline{\Psi}_1\ \Gamma^a\ \Psi_2)(\overline{\Psi}_3\ \Gamma^b\ \Psi_4)$ where a and b can taken the values: $\{\mathcal{S}, \mathcal{V}^\mu, \mathcal{T}^{\mu\nu}, \mathcal{P}, \mathcal{A}^\mu\}$ according as the bilinear that we are considering. We are required to find how this quantity can be expressed with the rearrangement, that is, say $(\overline{\Psi}_1\ \Gamma^a\ \Psi_4)(\overline{\Psi}_3\ \Gamma^b\ \Psi_2)$. For that, let us consider it as a linear combination, that is:

$$(\overline{\Psi}_1\ \Gamma_a\ \Psi_2)(\overline{\Psi}_3\ \Gamma^b\ \Psi_4) = \sum_{c,d}\mathcal{C}^{bc}_{ad}\ (\overline{\Psi}_1\ \Gamma_c\ \Psi_4)(\overline{\Psi}_3\ \Gamma^d\ \Psi_2)$$

Expanding the LHS of the above expression in terms of matrix indices, we get:

$$\sum_{\substack{i,j\\p,q}}\left\{(\overline{\Psi}_1)_i(\Gamma_a)_{ij}(\Psi_2)_j\right\}\left\{(\overline{\Psi}_3)_p(\Gamma^b)_{pq}(\Psi_4)_q\right\} = \sum_{\substack{i,j\\p,q}}(\overline{\Psi}_1)_i(\Gamma_a)_{ij}\left\{(\Psi_2)_j(\overline{\Psi}_3)_p\right\}\left(\Gamma^b\right)_{pq}(\Psi_4)_q$$

Now note that we can write:

$$(\Psi_2)_j(\overline{\Psi}_3)_p = \sum_{m,n}\delta_{jm}\delta_{pn}(\Psi_2)_m(\overline{\Psi}_3)_n$$

In Lec. 19.3, we had found a *completeness relation* in Eq. 13, for the elements of the basis and we can use this here.

$$(\Psi_2)_j(\overline{\Psi}_3)_p = \frac{1}{4}\sum_{m,n,c}(\Gamma_c)_{nm}(\Gamma^c)_{jp}(\Psi_2)_m(\overline{\Psi}_3)_n$$

Substituting this in the original expression, we have:

$$\begin{aligned}\text{LHS}: &\frac{1}{4}\sum_{\substack{m,n,c\\i,j,p,q}}(\overline{\Psi}_1)_i(\Gamma_a)_{ij}(\Gamma_c)_{nm}(\Gamma^c)_{jp}(\Psi_2)_m(\overline{\Psi}_3)_n(\Gamma^b)_{pq}(\Psi_4)_q \\ &= \frac{1}{4}\sum_{\substack{m,n,c\\i,j,p,q}}\left\{(\Gamma_c)_{nm}(\Psi_2)_m(\overline{\Psi}_3)_n\right\}(\overline{\Psi}_1)_i(\Gamma_a)_{ij}(\Gamma^c)_{jp}(\Gamma^b)_{pq}(\Psi_4)_q\end{aligned}$$

Note that the bracket term can be written as: $\text{Tr}\left(\Gamma_c\Psi_2\overline{\Psi}_3\right) = \overline{\Psi}_3\Gamma_c\Psi_2$ and the remaining term is written as:

$$\overline{\Psi}_1(\Gamma_a\Gamma^c\Gamma^b)\Psi_4$$

Since $\Gamma_a\Gamma^c\Gamma^b$ is a matrix, from Eq. 12 it can be expanded in terms of the basis elements accordingly as:

$$\Gamma_a\Gamma^c\Gamma^b = \sum_d \frac{1}{4}\text{Tr}\left(\Gamma_d\Gamma_a\Gamma^c\Gamma^b\right)\Gamma^d$$

Substituting this in the expression for LHS, we have:

$$(\overline{\Psi}_1\,\Gamma_a\,\Psi_2)(\overline{\Psi}_3\,\Gamma^b\,\Psi_4) = \frac{1}{16}\sum_{c,d}(\overline{\Psi}_3\Gamma_c\Psi_2)\overline{\Psi}_1\,\text{Tr}\left(\Gamma_d\Gamma_a\Gamma^c\Gamma^b\right)\Gamma^d\Psi_4 = \frac{1}{16}\sum_{c,d}\text{Tr}\left(\Gamma_d\Gamma_a\Gamma^c\Gamma^b\right)(\overline{\Psi}_3\Gamma_c\Psi_2)\overline{\Psi}_1\Gamma^d\Psi_4$$

Renaming the dummy indices $c \leftrightarrow d$ and arranging things accordingly, we obtain:

$$(\overline{\Psi}_1\,\Gamma_a\,\Psi_2)(\overline{\Psi}_3\,\Gamma^b\,\Psi_4) = \frac{1}{16}\sum_{c,d}\text{Tr}\left(\Gamma^c\Gamma_a\Gamma_d\Gamma^b\right)(\overline{\Psi}_1\Gamma_c\Psi_4)(\overline{\Psi}_3\Gamma^d\Psi_2)$$

Now, comparing the original linear combination, we obtain the form of the coefficients:

$$\boxed{\mathcal{C}^{bc}_{ad} = \frac{1}{16}\text{Tr}\left(\Gamma_a\Gamma_d\Gamma^b\Gamma^c\right)}$$

Let us now see some examples of this technique.

- $a = b = \mathcal{S}$ and so both are scalar bilinears, $\Gamma^a \equiv \mathbb{1}$ and $\Gamma_a \equiv \mathbb{1}$ Using the formula we have:

  $$\mathcal{C}^{\mathcal{S}\mathcal{S}}_{cd} = \frac{1}{16}\text{Tr}\left(\Gamma_c\Gamma^d\right) = \frac{1}{4}\delta^d{}_c$$

  Hence in this case, we can write the terms are:

  $$\begin{aligned}\overline{\Psi}_1\Psi_2\overline{\Psi}_3\Psi_4 = \frac{1}{4}\{&(\overline{\Psi}_1\Psi_4)(\overline{\Psi}_3\Psi_2) + (\overline{\Psi}_1\,\gamma_\mu\,\Psi_4)(\overline{\Psi}_3\,\gamma^\mu\,\Psi_2) + (\overline{\Psi}_1\,\gamma_5\,\Psi_4)(\overline{\Psi}_3\,\gamma^5\,\Psi_2)\\ &(\overline{\Psi}_1\,i\gamma_\mu\gamma_5\,\Psi_4)(\overline{\Psi}_3\,i\gamma^\mu\gamma_5\,\Psi_2) + \frac{1}{2}(\overline{\Psi}_1\,\Sigma_{\mu\nu}\,\Psi_4)(\overline{\Psi}_3\,\Sigma^{\mu\nu}\,\Psi_2)\}\end{aligned}$$

  The factor of $1/2$ has been introduced to avoid double counting since $\Sigma^{\mu\nu}$ has two indices. Thus, we find the rearranged spinors as:

  $$\boxed{\mathcal{S}_{12}\mathcal{S}_{34} = \frac{1}{4}\mathcal{S}_{14}\mathcal{S}_{32} + \frac{1}{4}(\mathcal{V}_\mu)_{14}(\mathcal{V}^\mu)_{32} + \frac{1}{4}\mathcal{P}_{14}\mathcal{P}_{32} - \frac{1}{4}(\mathcal{A}_\mu)_{14}(\mathcal{A}^\mu)_{32} + \frac{1}{8}(\mathcal{T}_{\mu\nu})_{14}(\mathcal{T}^{\mu\nu})_{32}}$$

- $a = b = \mathcal{V}$ and so both are vector bilinears. $\Gamma^a \equiv \gamma^\mu$ and $\Gamma_a \equiv \gamma_\mu = \eta_{\mu\nu}\gamma^\mu$. So the coefficient that we need to find out becomes: $\mathrm{Tr}(\gamma_\mu \Gamma_d \gamma^\mu \Gamma^c)$ Let us go case by case:

  $\Gamma_c = 1$

$$\begin{aligned} \frac{1}{16}\sum_c \mathrm{Tr}(\gamma_\mu \Gamma_d \gamma^\mu)(\bar{\Psi}_1\Psi_4)(\bar{\Psi}_3\Gamma^d\Psi_2) &= \frac{1}{16}\left[\bar{\Psi}_3 \sum_c \mathrm{Tr}(\gamma^\mu\gamma_\mu\Gamma_d)\Gamma^d\ \Psi_2\right](\bar{\Psi}_1\Psi_4) \\ &= \frac{1}{16}\times \bar{\Psi}_3(4\underbrace{\gamma^\mu\gamma_\mu}_{4\mathbb{1}})\Psi_2(\bar{\Psi}_1\Psi_4) \\ &= (\bar{\Psi}_1\Psi_4)(\bar{\Psi}_3\Psi_2) \to \mathcal{S}_{14}\mathcal{S}_{32} \end{aligned}$$

  $\Gamma_c = \gamma_\nu$

$$\begin{aligned} \frac{1}{16}\sum_c \mathrm{Tr}(\gamma_\mu \Gamma_d \gamma^\mu\gamma^\nu)(\bar{\Psi}_1\gamma_\nu\Psi_4)(\bar{\Psi}_3\Gamma^d\Psi_2) &= \frac{1}{16}\left[\bar{\Psi}_3 \sum_c \mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma_\mu\Gamma_d)\Gamma^d\ \Psi_2\right](\bar{\Psi}_1\gamma_\nu\Psi_4) \\ &= \frac{1}{16}\times \bar{\Psi}_3(4\underbrace{\gamma^\mu\gamma^\nu\gamma_\mu}_{-2\gamma^\nu})\Psi_2(\bar{\Psi}_1\gamma_\nu\Psi_4) \\ &= -\frac{1}{2}(\bar{\Psi}_1\gamma_\nu\Psi_4)(\bar{\Psi}_3\gamma^\nu\Psi_2) \to -\frac{1}{2}(\mathcal{V}_\nu)_{14}(\mathcal{V}^\nu)_{32} \end{aligned}$$

  $\Gamma_c = \gamma_5$

$$\begin{aligned} \frac{1}{16}\sum_c \mathrm{Tr}\big(\gamma_\mu \Gamma_d \gamma^\mu\gamma^5\big)(\bar{\Psi}_1\gamma_5\Psi_4)(\bar{\Psi}_3\Gamma^d\Psi_2) &= \frac{1}{16}\left[\bar{\Psi}_3 \sum_c \mathrm{Tr}\big(\gamma^\mu\gamma^5\gamma_\mu\Gamma_d\big)\Gamma^d\ \Psi_2\right](\bar{\Psi}_1\gamma_5\Psi_4) \\ &= \frac{1}{16}\times \bar{\Psi}_3(4\underbrace{\gamma^\mu\gamma^5\gamma_\mu}_{-4\gamma^5})\Psi_2(\bar{\Psi}_1\gamma_5\Psi_4) \\ &= -(\bar{\Psi}_1\gamma_5\Psi_4)(\bar{\Psi}_3\gamma^5\Psi_2) \to -\mathcal{P}_{14}\mathcal{P}_{32} \end{aligned}$$

  $\Gamma_c = i\gamma_\nu\gamma_5$

$$\begin{aligned} \frac{1}{16}\sum_c \mathrm{Tr}\big(\gamma_\mu \Gamma_d \gamma^\mu\ i\gamma^\nu\gamma^5\big)(\bar{\Psi}_1\ i\gamma_\nu\gamma_5\Psi_4)(\bar{\Psi}_3\Gamma^d\Psi_2) &= -\frac{1}{16}\left[\bar{\Psi}_3 \sum_c \mathrm{Tr}\big(\gamma^\mu\gamma^\nu\gamma^5\gamma_\mu\Gamma_d\big)\Gamma^d\ \Psi_2\right](\bar{\Psi}_1\gamma_\nu\gamma_5\Psi_4) \\ &= -\frac{1}{16}\times \bar{\Psi}_3(4\underbrace{\gamma^\mu\gamma^\nu\gamma^5\gamma_\mu}_{+2\gamma^\nu\gamma^5})\Psi_2(\bar{\Psi}_1\gamma_\nu\gamma_5\Psi_4) \\ &= -\frac{1}{2}(\bar{\Psi}_1\gamma_\nu\gamma_5\Psi_4)(\bar{\Psi}_3\gamma^\nu\gamma^5\Psi_2) \to -\frac{1}{2}(\mathcal{A}_\nu)_{14}(\mathcal{A}^\nu)_{32} \end{aligned}$$

  $\Gamma_c = \Sigma_{\rho\nu}$

$$\begin{aligned} \frac{1}{16}\sum_c \mathrm{Tr}(\gamma_\mu \Gamma_d \gamma^\mu\Sigma^{\rho\nu})(\bar{\Psi}_1\Sigma_{\rho\nu}\Psi_4)(\bar{\Psi}_3\Gamma^d\Psi_2) &= \frac{1}{16}\left[\bar{\Psi}_3 \sum_c \mathrm{Tr}(\gamma^\mu\Sigma^{\rho\nu}\gamma_\mu\Gamma_d)\Gamma^d\ \Psi_2\right](\bar{\Psi}_1\Sigma_{\rho\nu}\Psi_4) \\ &= \frac{1}{16}\times \bar{\Psi}_3(4\underbrace{\gamma^\mu\Sigma^{\rho\nu}\gamma_\mu}_{4(\eta^{\rho\nu}-\eta^{\nu\rho})})\Psi_2(\bar{\Psi}_1\Sigma_{\rho\nu}\Psi_4) = 0 \end{aligned}$$

Thus we have the final form:

$$\boxed{\mathcal{V}_{12}\mathcal{V}_{34} = 1\times\mathcal{S}_{14}\mathcal{S}_{32} + -\frac{1}{2}(\mathcal{V}_\mu)_{14}(\mathcal{V}^\mu)_{32} + -1\times\mathcal{P}_{14}\mathcal{P}_{32} - \frac{1}{2}(\mathcal{A}_\mu)_{14}(\mathcal{A}^\mu)_{32} + 0\times(\mathcal{T}_{\mu\nu})_{14}(\mathcal{T}^{\mu\nu})_{32}}$$

We can continue doing these kind of bs calculations and find more and more rearranged bilinears. However it must be noted that this is a very retarded task to do and one should never pursue this if it is not absolutely required!

	$\mathcal{S}_{14}\mathcal{S}_{32}$	$(\mathcal{V}_{14})_\nu(\mathcal{V}_{32})^\nu$	$(\mathcal{T}_{\rho\nu})_{14}(\mathcal{T}^{\rho\nu})_{32}$	$(\mathcal{A}_\nu)_{14}(\mathcal{A}^\nu)_{32}$	$\mathcal{P}_{14}\mathcal{P}_{32}$
$\mathcal{S}_{12}\mathcal{S}_{34}$	1/4	1/4	1/8	-1/4	1/4
$(\mathcal{V}_\mu)_{12}\mathcal{V}^\mu{}_{34}$	1	-1/2	0	-1/2	-1

Table 1: Table showing the coefficients of the rearranged spinors

Lecture 22: Normalising Spinors

Recall our discussion from Lec 17 where we derived the forms of the Dirac spinor for a momentum p (Eq. 9). This led to positive and negative energy states whose interpretation became problematic. The hole theory itself provided many inconsistencies and now another interpretation is followed. Let the proposed plane-wave solution to the Dirac equation be:

$$\Psi(x) = u(p)e^{-ipx} \qquad \Psi(x) = v(p)e^{+ipx}$$

where the $-$ is interpreted for *particles* and $+$ is interpreted for *anti-particles*. Note that in this case, $E > 0$ always. The distinction between particles and anti-particles comes from the sign of the exponential and hence we do not need to worry separately about negative and positive energy solutions.

We will define the following notation: $\not{a} := \gamma^\mu a_\mu$ henceforth to denote the contraction with the Dirac matrices. Using this, the Dirac equation is written as: $(i\not{\partial} - m)\Psi = 0$. Substituting the above solutions in the Dirac equation, we get:

$$\begin{aligned} 0 &= (i\not{\partial} - m)u(p)e^{-ip_\mu x^\mu} = i\gamma^\mu(-ip_\mu)e^{-ip_\mu x^\mu}u(p) - me^{-ip_\mu x^\mu}u(p) = (\not{p} - m)u(p)^{-ip_\mu x^\mu} \\ 0 &= (i\not{\partial} - m)v(p)e^{+ip_\mu x^\mu} = i\gamma^\mu(ip_\mu)e^{-ip_\mu x^\mu}v(p) - me^{ip_\mu x^\mu}v(p) = -(\not{p} + m)v(p)^{+ip_\mu x^\mu} \end{aligned}$$

From this, we get two separate equations that should be satisfied by the proposed solutions:

$$(\not{p} - m)u(p) = 0 \qquad (\not{p} + m)v(p) = 0 \tag{15}$$

In the Dirac basis the quantity $\not{p}$ becomes:

$$\not{p} = \gamma^\mu p_\mu = \gamma^0 p_0 + \gamma^i p_i = \gamma^0 p^0 - \gamma^i p^i = \begin{pmatrix} p^0\mathbb{1} & 0 \\ 0 & -p^0\mathbb{1} \end{pmatrix} - \begin{pmatrix} 0 & \sigma^i p^i \\ -\sigma^i p^i & 0 \end{pmatrix} = \begin{pmatrix} p^0\mathbb{1} & -\boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & -p^0\mathbb{1} \end{pmatrix} \tag{16}$$

Suppose the solutions be decomposed into a pair of two-component spinors such that:

$$u(p) = \begin{pmatrix} \xi(p) \\ \phi(p) \end{pmatrix} \qquad v(p) = \begin{pmatrix} \chi(p) \\ \eta(p) \end{pmatrix}$$

Substituting this in the above Eq. 15, we get the following:

$$\begin{aligned} (\boldsymbol{\sigma}\cdot\mathbf{p})\xi(p) - (p^0 + m)\phi(p) = 0 &\longrightarrow \phi(p) = \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0 + m}\xi(p) \\ (p^0 + m)\chi(p) - (\boldsymbol{\sigma}\cdot\mathbf{p})\eta(p) = 0 &\longrightarrow \chi(p) = \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0 + m}\eta(p) \end{aligned} \tag{17}$$

Thus the general solutions can be expressed in terms of two arbitrary Weyl spinors ξ and η. For simplicity, let us choose them to be the standard basis vectors of the space, that is,

$$\triangleright\ \xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \triangleright\ \xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \triangleright\ \eta^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \triangleright\ \eta^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The solutions can then be written as (for $s = 1, 2$):

$$u^s(p) = \begin{pmatrix} \xi^s \\ \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s \end{pmatrix} \qquad v^s(p) = \begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ \eta^s \end{pmatrix}$$

22.1. The correct factor

Let us try to find the correct normalisation for the spinors that we have obtained above.

$$u^{r\dagger}u^s = \begin{pmatrix}\xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\end{pmatrix}\begin{pmatrix}\xi^s \\ \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s\end{pmatrix} = \xi^{r\dagger}\xi^s + \frac{1}{(p^0+m)^2}\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p})^2\xi^s = \delta^{rs} + \frac{|\mathbf{p}|^2}{(p^0+m)^2}\delta^{rs}$$

We have use the following facts:

- ▷ The Pauli matrices are Hermitian, hence $(\boldsymbol{\sigma}\cdot\mathbf{p})^\dagger = (\boldsymbol{\sigma}\cdot\mathbf{p})$.
- ▷ $(AB)^\dagger = B^\dagger A^\dagger$.
- ▷ $(\boldsymbol{\sigma}\cdot\mathbf{p}) = |\mathbf{p}|^2\mathbb{1}$ (calculated in Eq. 8)
- ▷ $\xi^{r\dagger}\xi^s = \delta^{rs}$ since these are treated as the standard orthonormal basis vectors.

We know that $|\mathbf{p}|^2 = (p^0)^2 - m^2$ and substituting this above, we obtain:

$$u^{r\dagger}u^s = \left(1 + \frac{(p^0)^2 - m^2}{(p^0+m)^2}\right)\delta^{rs} = \frac{2p^0}{p^0+m}\delta^{rs} \tag{18}$$

Exactly same calculations yields for the other solution,

$$v^{r\dagger}v^s = \left(1 + \frac{(p^0)^2 - m^2}{(p^0+m)^2}\right)\delta^{rs} = \frac{2p^0}{p^0+m}\delta^{rs} \tag{19}$$

These are unfortunately not Lorentz invariant and we need to fix some factors to get a Lorentz invariant inner product between the spinors. Okay, so now let's check the orthogonality of the positive and negative frequency solutions.

$$u^{r\dagger}v^s = \begin{pmatrix}\xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\end{pmatrix}\begin{pmatrix}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ \eta^s\end{pmatrix} = \frac{2}{p^0+m}\xi^{r\dagger}\ (\boldsymbol{\sigma}\cdot\mathbf{p})\ \eta^s \neq 0 \tag{20}$$

The positive and negative frequency solutions are not orthogonal which is problematic since we expect the solutions for particles and anti-particles to not have any overlap.

We need to find some alternate definition of the inner product. Since the answer has already been found by major stalwarts, we just throw it in here. Consider the inner product $\bar{u}^r v^s = u^{r\dagger}\gamma^0 v^s$

$$u^{r\dagger}\gamma^0 v^s = \begin{pmatrix}\xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\end{pmatrix}\begin{pmatrix}\mathbb{1} & \\ & -\mathbb{1}\end{pmatrix}\begin{pmatrix}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ \eta^s\end{pmatrix} = \begin{pmatrix}\xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\end{pmatrix}\begin{pmatrix}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ -\eta^s\end{pmatrix} = 0$$

This definition of the inner product restores the *orthogonality* that we had desperately wanted! Let us look at the other inner products, that is, between the same sectors:

$$\begin{aligned}\bar{u}^r u^s = u^{r\dagger}\gamma^0 u^s = \begin{pmatrix}\xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\end{pmatrix}\begin{pmatrix}\mathbb{1} & \\ & -\mathbb{1}\end{pmatrix}\begin{pmatrix}\xi^s \\ \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s\end{pmatrix} &= \xi^{r\dagger}\xi^s - \frac{1}{(p^0+m)^2}\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p})^2\xi^s \\ &= \delta^{rs} - \frac{|\mathbf{p}|^2}{(p^0+m)^2}\delta^{rs} \\ &= \frac{2m}{p^0+m}\delta^{rs}\end{aligned}$$

Similar calculation yields, for the negative frequency solution,

$$\bar{v}^r v^s = -\frac{2m}{p^0+m}\delta^{rs}$$

All things look good except the factor $p^0 + m$ in the denominator which keeps these inner products from being Lorentz invariant. Hence if we impose the normalisation as:

$$u^s(p) = \sqrt{p^0+m}\begin{pmatrix} \xi^s \\ \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s \end{pmatrix} \qquad v^s(p) = \sqrt{p^0+m}\begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ \eta^s \end{pmatrix}$$

then, we will obtain $\bar{v}^r v^s = -2m\ \delta^{rs}$ and $\bar{u}^r u^s = 2m\ \delta^{rs}$ which is perfectly Lorentz invariant since it involved only the mass m which is a scalar. Henceforth, we will use this normalisation everytime we write the Dirac spinors. With this normalisation then, we get:

$$\boxed{u^{r\dagger}u^s = 2p^0\delta^{rs}} \qquad \boxed{v^{r\dagger}v^s = 2p^0\delta^{rs}} \tag{21}$$

Also note that, from Eq. 20, if we calculate the expression for $\mathbf{p}$ and $-\mathbf{p}$, we get:

$$\boxed{u^{r\dagger}(\mathbf{p})v^s(-\mathbf{p}) = u^{r\dagger}(-\mathbf{p})v^s(\mathbf{p}) = 0} \tag{22}$$

22.2. Outer Products

We looked at inner products and arrived at a nice definition of inner product for Dirac spinors. We now look at the outer product of these spinors.

$$\begin{aligned}
u^s\bar{u}^r = u^s u^{r\dagger}\gamma^0 &= \underbrace{(p^0+m)}_{\text{from normalisation}}\begin{pmatrix} \xi^s \\ \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s \end{pmatrix}\begin{pmatrix} \xi^{r\dagger} & \xi^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m} \end{pmatrix}\begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} (p^0+m)\xi^s\xi^{r\dagger} & \xi^s\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p})\xi^s\xi^{r\dagger} & \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) \end{pmatrix}\begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} (p^0+m)\xi^s\xi^{r\dagger} & -\xi^s\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p})\xi^s\xi^{r\dagger} & -\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\xi^s\xi^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) \end{pmatrix}
\end{aligned}$$

Similar for the other solution, we obtain the following form:

$$\begin{aligned}
v^s\bar{v}^r = v^s v^{r\dagger}\gamma^0 &= \underbrace{(p^0+m)}_{\text{from normalisation}}\begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s \\ \eta^s \end{pmatrix}\begin{pmatrix} \eta^{r\dagger}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m} & \eta^{r\dagger} \end{pmatrix}\begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s\eta^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) & (\boldsymbol{\sigma}\cdot\mathbf{p})\eta^s\eta^{r\dagger} \\ \eta^s\eta^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) & (p^0+m)\eta^s\eta^{r\dagger} \end{pmatrix}\begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}\eta^s\eta^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) & -(\boldsymbol{\sigma}\cdot\mathbf{p})\eta^s\eta^{r\dagger} \\ \eta^s\eta^{r\dagger}(\boldsymbol{\sigma}\cdot\mathbf{p}) & -(p^0+m)\eta^s\eta^{r\dagger} \end{pmatrix}
\end{aligned}$$

As ξ and η are basis vectors, these satisfy the completeness relation:

$$\sum_r \eta^r\eta^{r\dagger} = \mathbb{1} \qquad \sum_r \xi^r\xi^{r\dagger} = \mathbb{1}$$

Using this, we get:

$$\sum_r u^r \bar{u}^r = \begin{pmatrix} (p^0+m)\mathbb{1} & -(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}(\boldsymbol{\sigma}\cdot\mathbf{p}) \end{pmatrix} = \begin{pmatrix} (p^0+m)\mathbb{1} & -(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -\frac{|\mathbf{p}|^2}{p^0+m}\mathbb{1} \end{pmatrix} = \begin{pmatrix} (p^0+m)\mathbb{1} & -(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -(p^0-m)\mathbb{1} \end{pmatrix} = \not{p} + m\mathbb{1}$$

$$\sum_r v^r \bar{v}^r = \begin{pmatrix} \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p^0+m}(\boldsymbol{\sigma}\cdot\mathbf{p}) & -(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -(p^0+m)\mathbb{1} \end{pmatrix} = \begin{pmatrix} (p^0-m)\mathbb{1} & -(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -(p^0+m)\mathbb{1} \end{pmatrix} = \not{p} - m\mathbb{1}$$

Thus we have two nice identitites concerning the outer products,

$$\boxed{\sum_r u^r \bar{u}^r = \not{p} + m} \qquad \boxed{\sum_r v^r \bar{v}^r = \not{p} - m} \tag{23}$$

We had used the Dirac representation to find these identitites. We could have also used the representation of Dirac matrices in chiral basis which would have given us the same result, since the final expression is independent of the representation of the Dirac matrix.

22.3. Projecting Energy

Let us define the following operator:

$$\Lambda_\pm := \frac{\pm\not{p} + m}{2m}$$

Let us see some properties of these operators:

▷ Idempotence

$$\Lambda_+^2 = \frac{1}{4m^2}(\not{p}+m)(\not{p}+m) = \frac{1}{4m^2}(\not{p}\not{p} + 2m\not{p} + m^2) = \frac{1}{4m^2}(2m^2 + 2m\not{p}) = \Lambda_+$$

$$\Lambda_-^2 = \frac{1}{4m^2}(-\not{p}+m)(-\not{p}+m) = \frac{1}{4m^2}(\not{p}\not{p} - 2m\not{p} + m^2) = \frac{1}{4m^2}(2m^2 - 2m\not{p}) = \Lambda_-$$

where we have used the following:

$$\not{p}\not{p} = \frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)p_\mu p_\nu = \eta^{\mu\nu}p_\mu p_\nu = p^2 = m^2$$

Thus, we find that $\Lambda_\pm$ are idempotent operators.

▷ Orthogonality

$$\Lambda_+\Lambda_- \sim (\not{p}+m)(-\not{p}+m) = (-m^2+m^2) = 0$$

$$\Lambda_-\Lambda_+ \sim (-\not{p}+m)(\not{p}+m) = (-m^2+m^2) = 0$$

Thus, we see that product of the two operators always give zero.

▷ Completeness

$$\Lambda_+ + \Lambda_- = \frac{1}{2m}(\not{p} + m - \not{p} + m) = \mathbb{1}$$

We see that the sum of the operators equal to the identity operator of the space.

The above properties indicate that $\Lambda_\pm$ are *orthogonal projection operators*. To see explicitly, consider the following:

$$\Lambda_+ u^s = \frac{(\not{p}+m)}{2m}u^s = \frac{mu^s + mu^s}{2m} = u^s$$

$$\Lambda_- u^s = \frac{(-\not{p}+m)}{2m}u^s = \frac{-mu^s + mu^s}{2m} = 0$$

$$\Lambda_+ v^s = \frac{(\not{p}+m)}{2m}v^s = \frac{-mv^s + mv^s}{2m} = 0$$

$$\Lambda_+ v^s = \frac{(-\not{p}+m)}{2m}v^s = \frac{mv^s + mv^s}{2m} = v^s$$

where we have used Eq. 15. We see that, starting from an arbitrary spinor, Λ_+ (Λ_-) projects to the positive (negative) frequency sector. Hence these are also called *energy projection operators*.

Lecture 23: Some Other Projections

23.1. Spin Projection

In this section we will be concerned with spins. For the *non-relativistic* case, the *spin projection operator* is given by:

$$\Sigma(\hat{s}) = \frac{1}{2}(1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{s}})$$

Think about it, suppose $\hat{\mathbf{s}} = (0, 0, \pm 1)$ then essentially we have a spin state either in $+z$ or $-z$ direction, denoted by say, $|\alpha\rangle$. The operator then becomes:

$$\Sigma_z = \frac{1}{2}(1 \pm \sigma_z)\,|\alpha\rangle = \frac{1}{2}(1 \pm s)\,|\alpha\rangle$$

where $s = \pm 1$ is the eigenvalue of σ_z. If $s = +1$ and $\hat{\mathbf{s}} = (0, 0, 1)$ or $s = -1$ and $\hat{\mathbf{s}} = (0, 0, -1)$, then $\Sigma_z = |\alpha\rangle$, else $\Sigma_z = 0$. This correctly projects an arbitrary spinor to the appropriate spin $\hat{\mathbf{s}}$.

We need to extrapolate this idea to the relativistic case. First, let us begin considering the REST FRAME, where $p^\mu = (p^0, 0)$. Consider the operator,

$$\Sigma(\hat{s}) := \frac{1}{2}\begin{pmatrix} 1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{s}} & 0 \\ 0 & 1 - \boldsymbol{\sigma} \cdot \hat{\mathbf{s}} \end{pmatrix} = \frac{1}{2}\left[\mathbb{1} + \begin{pmatrix} \boldsymbol{\sigma} \cdot \hat{\mathbf{s}} & 0 \\ 0 & -\boldsymbol{\sigma} \cdot \hat{\mathbf{s}} \end{pmatrix}\right]$$

We define the quantity $\hat{s}^\mu := (0, \hat{\mathbf{s}})$. Then, using the same calculation as in Eq.16, we have:

$$\not{\hat{s}} = \begin{pmatrix} 0 & -\boldsymbol{\sigma} \cdot \hat{\mathbf{s}} \\ \boldsymbol{\sigma} \cdot \hat{\mathbf{s}} & 0 \end{pmatrix}$$

since $\hat{s}^0 = 0$. Then $\Sigma(\hat{s})$ can be written as $\Sigma(\hat{s}) = \frac{1}{2}\left[\mathbb{1} + \gamma^5 \not{\hat{s}}\right]$. Also, note a few properties:

$$\hat{s}^\mu \hat{s}_\mu = -|\hat{\mathbf{s}}|^2 = -1 \qquad p^\mu s_\mu = 0 \tag{24}$$

Let us see what happens when we act this to the free-wave solutions of the Dirac equation, $u^r(p)$ and $v^r(p)$. Then the solutions become:

$$u^r = \begin{pmatrix} \xi^r \\ 0 \end{pmatrix} \qquad v^r = \begin{pmatrix} 0 \\ \eta^r \end{pmatrix} \longrightarrow \Sigma(\hat{s})u^r = \frac{1}{2}\begin{pmatrix} (1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{s}})\xi^r \\ 0 \end{pmatrix} \qquad \Sigma(\hat{s})v^r = \frac{1}{2}\begin{pmatrix} 0 \\ (1 - \boldsymbol{\sigma} \cdot \hat{\mathbf{s}})\eta^r \end{pmatrix}$$

For simplicity, let us take $\hat{\mathbf{s}} = (0, 0, \pm 1)$ (that is, spin is along $\hat{z}$ axis). We can have an interpretation of u^r and v^r with r indicating the spin, that is, $r = 1$ corresponds to spin $+\frac{1}{2}$ and $r = 2$ corresponds to spin $-\frac{1}{2}$ for the positive energy solution u while the opposite occurs for v.

For $s = \pm 1$, it is now better to denote these solutions as $u(0, s)$ and $v(0, s)$ which mentions the spin explicitly. Hence $u^1 = u(0, +1), u^2 = u(0, -1), v^1 = v(0, -1), v^2 = v(0, +1)$ With this notation, one can verify that:

- If $\hat{\mathbf{s}} = (0, 0, 1)$ then $\Sigma(\hat{s})u^1 = u^1$
- If $\hat{\mathbf{s}} = (0, 0, -1)$ then $\Sigma(\hat{s})u^1 = 0$
- If $\hat{\mathbf{s}} = (0, 0, 1)$ then $\Sigma(\hat{s})u^2 = 0$
- If $\hat{\mathbf{s}} = (0, 0, -1)$ then $\Sigma(\hat{s})u^2 = u^2$
- If $\hat{\mathbf{s}} = (0, 0, 1)$ then $\Sigma(\hat{s})v^1 = 0$
- If $\hat{\mathbf{s}} = (0, 0, -1)$ then $\Sigma(\hat{s})v^1 = v^1$
- If $\hat{\mathbf{s}} = (0, 0, 1)$ then $\Sigma(\hat{s})v^2 = v^2$
- If $\hat{\mathbf{s}} = (0, 0, -1)$ then $\Sigma(\hat{s})v^2 = 0$

We see that in the rest frame, the free-wave solutions are eigenstates of $\Sigma(\hat{s})$ and they correctly project the right spin states, that is,

$$\Sigma(\hat{s})u(0,\hat{s}) = u(0,\hat{s})$$

The equation above has a invariant form, that is, it has no index dependence. We can use this now to go into a general frame by considering $p \to \Lambda(m,0)$ and $\hat{s} \to \Lambda\hat{s}$. Then note that,

$$\begin{aligned} u(p,s) = S(\Lambda)u(0,\hat{s}) &= S(\Lambda)\ \Sigma(\hat{s})\ S^{-1}(\Lambda)S(\Lambda)\ u(0,\hat{s}) \\ &= \frac{1}{2}\big[\mathbb{1} + \gamma^5 S(\Lambda)\ \gamma^\mu\ S^{-1}(\Lambda)\hat{s}_\mu\big]S(\Lambda)u(0,\hat{s}) \\ &= \frac{1}{2}\Big[\mathbb{1} + \gamma^5 \big(\Lambda^{-1}\big)^\mu{}_\nu \gamma^\nu \hat{s}_\mu\Big]u(p,s) \\ &= \frac{1}{2}\big[\mathbb{1} + \gamma^5\gamma^\nu s_\nu\big]u(p,s) \end{aligned}$$

Hence in a general frame, the spin projection operator takes the form:

$$\Sigma(\pm s) = \frac{1}{2}\big[1 \pm \gamma^5 \not{s}\big]$$

where s^μ is now any vector satisfying Eq.24. We can check that these are indeed projection operators. For that, first let us calculate some quantitites which appear in the expression:

$$\gamma^\mu\gamma^\nu s_\mu s_\nu = (2\eta^{\mu\nu} - \gamma^\nu\gamma^\mu)s_\mu s_\nu = 2s^\nu s_\nu - \gamma^\nu\gamma^\mu s_\mu s_\nu = -2 - \gamma^\mu\gamma^\nu s_\nu s_\mu = -2 - \gamma^\mu\gamma^\nu s_\mu s_\nu \implies \boxed{\gamma^\mu\gamma^\nu s_\mu s_\nu = -1}$$

And we also have the following thing using the above identity:

$$\gamma^5\not{s}\gamma^5\not{s} = \gamma^5\gamma^\mu\gamma^5\gamma^\nu s_\mu s_\nu = -(\gamma^5)^2\gamma^\mu\gamma^\nu s_\mu s_\nu = 1$$

Now, we are in a good position to calculate the projection properties.

- $\Sigma(s)\Sigma(s) = \frac{1}{4}(1 + 2\gamma^5\not{s} + \gamma^5\not{s}\gamma^5\not{s}) = \frac{1}{4}(1 + 2\gamma^5\not{s} + 1) = \Sigma(s)$
- $\Sigma(-s)\Sigma(-s) = \frac{1}{4}(1 - 2\gamma^5\not{s} + \gamma^5\not{s}\gamma^5\not{s}) = \frac{1}{4}(1 - 2\gamma^5\not{s} + 1) = \Sigma(-s)$
- $\Sigma(s)\Sigma(-s) = \frac{1}{4}(1 - \gamma^5\not{s}\gamma^5\not{s}) = 0$
- $\Sigma(s) + \Sigma(-s) = \frac{1}{2}(\mathbb{1} + \gamma^5\not{s} + \mathbb{1} - \gamma^5\ \not{s}) = 0$

We see that $\Sigma(\pm s)$ satisfies all the properties of orthogonal projection and is thus a valid projection operator.

In the previous section, we had seen that the energy projection operators $\Lambda_\pm$ projected states to a particular energy and here we found two operators projecting out to a particular spin state. Then we can project to some state with definite energy and definite spin by successivly acting the energy and spin projection operators together. However, problem might arise if these do not commute since then the answer would depend on their commutator. Hence, let us first see if these commute or not:

$$[\Lambda_\pm, \Sigma(\pm s)] \propto \big[\not{p}, \gamma^5\not{s}\big] = (\gamma^\mu p_\mu)(\gamma^5\gamma^\nu s_\nu) - (\gamma^5\gamma^\nu s_\nu)(\gamma^\mu p_\mu) = -\gamma^5(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)p_\mu s_\nu = -2\eta^{\mu\nu}p_\mu s_\nu = 0$$

where we have used Eq.24. We see that the energy projection and spin projection operators happily commute and we are out of trouble!

23.2. Chirality and Massless Fermions

We will focus on *chirality* now which is a Lorentz invariant property (however, unlike helicity it does not commute with the Hamiltonian). γ^5, having eigenvalues 1 and -1, is called the chirality operator and we define the right and left chiral states as the eigenstates of γ^5, that is:

$$\gamma^5 u_{\mathrm{R}} = u_{\mathrm{R}} \qquad \gamma^5 u_{\mathrm{L}} = -u_{\mathrm{L}}$$

From this definition, we can guess a form of the chirality projection operator:

$$P_{\mathrm{R}} = \frac{1}{2}(\mathbb{1} + \gamma^5) \qquad P_{\mathrm{L}} = \frac{1}{2}(\mathbb{1} - \gamma^5)$$

We can check that

$$\triangleright\ P_{\mathrm{R}} u_{\mathrm{R}} = u_{\mathrm{R}} \qquad \triangleright\ P_{\mathrm{R}} u_{\mathrm{L}} = 0 \qquad \triangleright\ P_{\mathrm{L}} u_{\mathrm{R}} = 0 \qquad \triangleright\ P_{\mathrm{L}} u_{\mathrm{L}} = u_{\mathrm{L}}$$

Also we can check the projection properties:

- $\frac{1}{4}(\mathbb{1} + \gamma^5)(\mathbb{1} + \gamma^5) = \frac{1}{4}(\mathbb{1} + 2\gamma^5 + (\gamma^5)^2) = \frac{1}{2}(\mathbb{1} + \gamma^5)$
- $\frac{1}{4}(\mathbb{1} - \gamma^5)(\mathbb{1} - \gamma^5) = \frac{1}{4}(\mathbb{1} - 2\gamma^5 + (\gamma^5)^2) = \frac{1}{2}(\mathbb{1} - \gamma^5)$
- $(\mathbb{1} + \gamma^5)(\mathbb{1} - \gamma^5) = \mathbb{1} - (\gamma^5)^2 = 0$
- $\frac{1}{2}\big\{(\mathbb{1} + \gamma^5) + (\mathbb{1} - \gamma^5)\big\} = \mathbb{1}$

and thus, this is a valid projection operator. Any spinor can be written in terms of the chiral states as:

$$\Psi = \mathbb{1}\Psi = (P_{\mathrm{L}} + P_{\mathrm{R}})\Psi = \Psi_{\mathrm{L}} + \Psi_{\mathrm{R}} \qquad \text{where,} \quad \Psi_{\mathrm{L}} := P_{\mathrm{L}}\Psi \quad \Psi_{\mathrm{R}} := P_{\mathrm{R}}\Psi$$

In the chiral basis,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \implies \gamma^5 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \longrightarrow P_{\mathrm{R}} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix} \quad P_{\mathrm{L}} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$$

and we see that, if any arbitrary Dirac spinor Ψ is written in terms of the Weyl spinors as,

$$\Psi = \begin{pmatrix} \psi_{\mathrm{L}} \\ \psi_{\mathrm{R}} \end{pmatrix}$$

then, the chirality projections are respectively

$$\Psi_{\mathrm{R}} = P_{\mathrm{R}}\Psi = \begin{pmatrix} 0 \\ \psi_{\mathrm{R}} \end{pmatrix} \qquad \Psi_{\mathrm{L}} = P_{\mathrm{L}}\Psi = \begin{pmatrix} \psi_{\mathrm{L}} \\ 0 \end{pmatrix}$$

and we were right to name the two-component spinors as *left* and handed Weyl spinors in the chiral basis. However, in the Dirac basis,

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \implies \gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \longrightarrow P_{\mathrm{R}} = \frac{1}{2}\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} \quad P_{\mathrm{L}} = \frac{1}{2}\begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}$$

Hence for any Dirac spinor written in terms of some arbitrary two component spinors ϕ and χ, we have:

$$\Psi_{\mathrm{R}} = P_{\mathrm{R}}\Psi = \frac{1}{2}\begin{pmatrix} \phi + \chi \\ \phi + \chi \end{pmatrix} \qquad \Psi_{\mathrm{L}} = P_{\mathrm{L}}\Psi = \frac{1}{2}\begin{pmatrix} \phi - \chi \\ \chi - \phi \end{pmatrix}$$

Then, if we define $\omega_{\mathrm{R}} = \frac{1}{2}(\phi + \chi)$ and $\omega_{\mathrm{L}} = \frac{1}{2}(\phi - \chi)$, then we get:

$$\Psi_{\mathrm{R}} = \begin{pmatrix} \omega_{\mathrm{R}} \\ \omega_{\mathrm{R}} \end{pmatrix} \qquad \Psi_{\mathrm{L}} = \begin{pmatrix} \omega_{\mathrm{L}} \\ -\omega_{\mathrm{L}} \end{pmatrix} \tag{25}$$

NOTE: ϕ and χ are not Weyl spinors, these are some arbitrary spinors. Above, Ψ_{L} and Ψ_{R} are eigenstates of γ^5 (as can be checked since in Dirac basis, the eigenstates of γ^5 are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$) and thus, have definite chirality. So, the two component spinors with which these are made up of, that is, ω_{L} and ω_{R} are the correct Weyl spinors.

23.2.1. When mass vanishes!

Suppose we have a Dirac spinor in terms of two-component spinors in the Dirac basis, $\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$.

Then using Eq. 16 and taking $m = 0$ we have:

$$\left.\begin{aligned} p^0\psi_1 - (\boldsymbol{\sigma}\cdot\mathbf{p})\psi_2 &= 0 \\ p^0\psi_2 - (\boldsymbol{\sigma}\cdot\mathbf{p})\psi_1 &= 0 \end{aligned}\right\} \qquad \boxed{p^0(\psi_1+\psi_2) = (\boldsymbol{\sigma}\cdot\mathbf{p})(\psi_1+\psi_2)} \quad \boxed{p^0(\psi_1-\psi_2) = -(\boldsymbol{\sigma}\cdot\mathbf{p})(\psi_1-\psi_2)} \tag{26}$$

Defining $\psi_{\mathrm{L}} = \frac{1}{2}(\psi_1 - \psi_2)$ and $\psi_{\mathrm{R}} = \frac{1}{2}(\psi_1 + \psi_2)$, and multiplying both sides of the equation with $(\boldsymbol{\sigma}\cdot\mathbf{p})$, we get:

$$p^0 \underbrace{(\boldsymbol{\sigma}\cdot\mathbf{p})\psi_{\mathrm{L}}}_{p^0\psi_{\mathrm{L}}} = (\boldsymbol{\sigma}\cdot\mathbf{p})^2\psi_{\mathrm{L}} = |\mathbf{p}|^2\psi_{\mathrm{L}} \longrightarrow (p^0)^2\psi_{\mathrm{L}} = |\mathbf{p}|^2\psi_{\mathrm{L}}$$
$$p^0 \underbrace{(\boldsymbol{\sigma}\cdot\mathbf{p})\psi_{\mathrm{R}}}_{-p^0\psi_{\mathrm{R}}} = -(\boldsymbol{\sigma}\cdot\mathbf{p})^2\psi_{\mathrm{L}} = -|\mathbf{p}|^2\psi_{\mathrm{L}} \longrightarrow (p^0)^2\psi_{\mathrm{R}} = |\mathbf{p}|^2\psi_{\mathrm{R}}$$

For a non-trivial solution, $p^0 = \pm|\mathbf{p}|$ and putting this in Eq. 26,

$$\left.\begin{aligned} \tfrac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{|\mathbf{p}|}\psi_{\mathrm{L}} &= \psi_{\mathrm{L}} \\ \tfrac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{|\mathbf{p}|}\psi_{\mathrm{R}} &= -\psi_{\mathrm{R}} \end{aligned}\right\} p^0 = +|\mathbf{p}| \qquad \left.\begin{aligned} \tfrac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{|\mathbf{p}|}\psi_{\mathrm{L}} &= -\psi_{\mathrm{L}} \\ \tfrac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{|\mathbf{p}|}\psi_{\mathrm{R}} &= \psi_{\mathrm{R}} \end{aligned}\right\} p^0 = -|\mathbf{p}| \tag{27}$$

Recall from Eq. 11 that the helicity operator was,

$$\hat{h} = \frac{\hbar}{2|\mathbf{p}|}(\boldsymbol{\Sigma}\cdot\mathbf{p}) = \frac{\hbar}{2|\mathbf{p}|}\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix}$$

Then, if we act $\Psi_{\mathrm{L}} = P_{\mathrm{L}}\Psi$ and $\Psi_{\mathrm{R}} = P_{\mathrm{R}}\Psi$ with the helicity operator (and using their form from Eq. 25), we have:

$$\hat{h}(P_{\mathrm{L}}\Psi) = \frac{\hbar}{2|\mathbf{p}|}\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix}\begin{pmatrix} \psi_{\mathrm{L}} \\ -\psi_{\mathrm{L}} \end{pmatrix} = \frac{\hbar}{2}(P_{\mathrm{L}}\Psi)$$
$$\hat{h}(P_{\mathrm{R}}\Psi) = \frac{\hbar}{2|\mathbf{p}|}\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix}\begin{pmatrix} \psi_{\mathrm{R}} \\ \psi_{\mathrm{R}} \end{pmatrix} = -\frac{\hbar}{2}(P_{\mathrm{L}}\Psi)$$

Hence for *particles* the left and right chiral states are also the helicity eigenstates with eigenvalue $\pm\hbar/2$ respectively (while for *anti-particles* the left and right chiral states are helicity eigenstates with eigenvalue $\mp\hbar/2$ which can be checked using the same calculation, hence chirality is negative of helicity for anti-particles).

For massive particles, it is possible to Lorentz boost to different frames of reference to change helicity (however chirality does not change since it is Lorentz invariant) and hence helicity and chirality are equivalent only for *massless* fermions.

Lecture 24: Quantisation of Dirac Field: What not to do!

Till now we have seen many different aspects of the Dirac equations and now we want to quantise it. For that, we first need to figure out a Lorentz invariant Lagrangian for the Dirac field.

24.1. Dirac Lagrangian

The appropriate Lagrangian turns out to be,

$$\boxed{\mathcal{L}_{\mathrm{D}} = \overline{\Psi}\ (i\partial\!\!\!/ - m)\ \Psi = \overline{\Psi}\ (i\gamma^\mu \partial_\mu - m)\ \Psi}$$

where Ψ and $\overline{\Psi}$ are now treated as two independent fields. Let us see whether this Lagrangian gives us the correct equation of motion or not!

$$0 = \partial_\mu\left(\frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_\mu \Psi)}\right) - \frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial \Psi} = i\partial_\mu(\overline{\Psi}\gamma^\mu) - (-m\overline{\Psi}) = i\partial_\mu\overline{\Psi}\gamma^\mu + m\overline{\Psi} = i\partial_\mu\Psi^\dagger\gamma^0\gamma^\mu + m\Psi^\dagger\gamma^0$$

We use the identity from Eq. 14 and the invertibility of γ^0 to further simplify this:

$$0 = i\partial_\mu\Psi^\dagger(\gamma^\mu)^\dagger\gamma^0 + m\Psi^\dagger\gamma^0 = (i\partial_\mu\Psi^\dagger(\gamma^\mu)^\dagger + m\Psi^\dagger)\gamma^0 \implies 0 = (i\partial_\mu(\gamma^\mu\Psi)^\dagger + m\Psi^\dagger) \implies \boxed{0 = (i\gamma^\mu\partial_\mu\Psi + m\Psi)}$$

We thus obtain the Dirac equation. Starting with $\overline{\Psi}$ we obtain in an instant:

$$0 = \partial_\mu\left(\frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_\mu \overline{\Psi})}\right) - \frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial \overline{\Psi}} = 0 - (i\partial\!\!\!/\Psi - m\Psi) \implies \boxed{(i\gamma^\mu\partial_\mu\Psi + m\Psi) = 0}$$

24.2. Conserved Current

The Dirac Lagrangian exhibits a global U(1) symmetry since if we consider

$$\Psi \longrightarrow e^{i\alpha}\Psi \implies \overline{\Psi} \longrightarrow e^{-i\alpha}\overline{\Psi}$$

Then,

$$\mathcal{L}_{\mathrm{D}} \longrightarrow \mathcal{L}'_{\mathrm{D}} = ie^{-i\alpha}\overline{\Psi}\gamma^\mu e^{i\alpha}\Psi - me^{-i\alpha}e^{i\alpha}\overline{\Psi}\Psi = \mathcal{L}$$

Under the infinitesimal transformation of the Dirac field,

$$\delta\Psi = \Psi' - \Psi = \left(e^{i\alpha} - 1\right)\Psi = [(1 + i\alpha) - 1]\Psi = i\alpha\Psi$$

using Eq. 1 we can find the conserved Noether current as,

$$J^\mu = \frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_\mu \Psi)}\delta\Psi + \cancelto{0}{\frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_\mu \overline{\Psi})}\delta\overline{\Psi}} = (i\overline{\Psi}\gamma^\mu)(i\alpha\Psi) = -\overline{\Psi}\gamma^\mu\Psi \tag{28}$$

24.3. Hamiltonian Density

Let us first find the conjugate momentum for the fields.

$$\pi_\Psi = \frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_0 \psi)} = i\overline{\Psi}\gamma^0 = i\Psi^\dagger(\gamma^0)^2 = i\Psi^\dagger \qquad \pi_{\overline{\Psi}} = \frac{\partial \mathcal{L}_{\mathrm{D}}}{\partial(\partial_0 \overline{\psi})} = 0$$

Then using the Legendre transform, we have the Hamiltonian density as:

$$\mathcal{H} = \pi_\Psi\dot{\Psi} + \pi_{\overline{\Psi}}\dot{\overline{\Psi}} - \mathcal{L} = i\overline{\Psi}\gamma^0\partial_0\Psi - i\overline{\Psi}\gamma^\mu\partial_\mu\Psi + m\overline{\Psi}\Psi = -i\overline{\Psi}\gamma^i\partial_i\Psi + m\overline{\Psi}\Psi$$

We had obtained the Lagrangian and the Hamiltonian density for the Dirac field. What we need to do now is to propose some nice guess for the field expressions. We had already found four solutions for a momentum p namely $u^s(p)$ and $v^s(p)$ for $s = 1, 2$. We can thus make the field a linear combination of these four solutions with some arbitrary coefficients:

$$\Psi(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(\hat{a}^s_{\mathbf{p}} u^s(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^s_{\mathbf{p}} v^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right)$$

Taking the Hermitian adjoint and multiplying by γ^0 on the right, we get the expansion for the Dirac adjoint,

$$\overline{\Psi}(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_r \left(\hat{a}^{r\dagger}_{\mathbf{p}} \overline{u}^r(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^{r\dagger}_{\mathbf{p}} \overline{v}^r(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right)$$

24.4. Total Hamiltonian

We will now try to find an expression for the total Hamiltonian $\hat{H}$:

$$\hat{H} = \int \mathrm{d}^3\mathbf{x}\ \mathcal{H} = \int \mathrm{d}^3\mathbf{x}\ \left(\overline{\Psi}(-i\gamma^i\partial_i + m)\Psi\right)$$

This is an extremely tedious and terrible calculation, one that I wouldn't wish upon anyone. Substituting the expression for $\Psi(\mathbf{x})$ and $\overline{\Psi}(\mathbf{x})$ we get:

$$\begin{aligned}\hat{H} = \int \mathrm{d}^3\mathbf{x} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_\mathbf{p}E_\mathbf{q}}}\sum_{r,s}\left(\hat{a}^{s\dagger}_\mathbf{p}u^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^{s\dagger}_\mathbf{p}v^s(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right)(-i\gamma^i\partial_i + m) \\ \times\left(\hat{a}^r_\mathbf{q}u^r(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} + \hat{b}^r_\mathbf{q}v^r(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{x}}\right)\end{aligned}$$

Note that $\mathbf{p}\cdot\mathbf{x} = x^ip^i = -x^ip_i$ and then $\partial_i \exp(\pm i\mathbf{p}\cdot\mathbf{x}) = \mp ip_i\exp(\pm i\mathbf{p}\cdot\mathbf{x})$. We use this in the above expression to get

$$\begin{aligned}\hat{H} = \int \mathrm{d}^3\mathbf{x} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_\mathbf{p}E_\mathbf{q}}}\sum_{r,s}\left(\hat{a}^{s\dagger}_\mathbf{p}\bar{u}^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^{s\dagger}_\mathbf{p}\bar{v}^s(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right) \\ \times\left(\hat{a}^r_\mathbf{q}(-\gamma^iq_i + m)u^r(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} + \hat{b}^r_\mathbf{q}(\gamma^iq_i + m)v^r(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{x}}\right)\end{aligned}$$

From Eq. 15, recall that we had the following:

$$\begin{aligned}(\not{p} - m)u^s(\mathbf{p}) = 0 &\implies (\gamma^0p_0 + \gamma^ip_i - m)u^s(\mathbf{p}) = 0 \implies (-\gamma^ip_i + m)u^s(\mathbf{p}) = E_\mathbf{p}\gamma^0u^s(\mathbf{p}) \\ (\not{p} + m)v^s(\mathbf{p}) = 0 &\implies (\gamma^0p_0 + \gamma^ip_i + m)v^s(\mathbf{p}) = 0 \implies (\gamma^ip_i + m)v^s(\mathbf{p}) = -E_\mathbf{p}\gamma^0v^s(\mathbf{p})\end{aligned}$$

Substituting this in the integral above we get:

$$\begin{aligned}\hat{H} &= \int \mathrm{d}^3\mathbf{x} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_\mathbf{p}E_\mathbf{q}}}E_\mathbf{q}\sum_{r,s}\left(\hat{a}^{s\dagger}_\mathbf{p}\bar{u}^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^{s\dagger}_\mathbf{p}\bar{v}^s(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right) \\ &\qquad\times\gamma^0\left(\hat{a}^r_\mathbf{q}u^r(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} - \hat{b}^r_\mathbf{q}v^r(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{x}}\right) \\ &= \int \mathrm{d}^3\mathbf{x} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\sqrt{\frac{E_\mathbf{q}}{4E_\mathbf{p}}}\sum_{r,s}\left(\hat{a}^{s\dagger}_\mathbf{p}u^{s\dagger}(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + \hat{b}^{s\dagger}_\mathbf{p}v^{s\dagger}(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right) \\ &\qquad\times(\gamma^0)^2\left(\hat{a}^r_\mathbf{q}u^r(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} - \hat{b}^r_\mathbf{q}v^r(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{x}}\right) \\ &= \int \mathrm{d}^3\mathbf{x} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\sqrt{\frac{E_\mathbf{q}}{4E_\mathbf{p}}}\sum_{r,s}(\hat{a}^{s\dagger}_\mathbf{p}\hat{a}^r_\mathbf{q}u^{s\dagger}(\mathbf{p})u^r(\mathbf{q})e^{i(\mathbf{q}-\mathbf{p})\cdot\mathbf{x}} - \hat{a}^{s\dagger}_\mathbf{p}\hat{b}^r_\mathbf{q}u^{s\dagger}(\mathbf{p})v^r(\mathbf{q})e^{-i(\mathbf{q}+\mathbf{p})\cdot\mathbf{x}} \\ &\qquad+\hat{b}^{s\dagger}_\mathbf{p}\hat{a}^r_\mathbf{q}v^{s\dagger}(\mathbf{p})u^r(\mathbf{q})e^{i(\mathbf{q}+\mathbf{p})\cdot\mathbf{x}} - \hat{b}^{s\dagger}_\mathbf{p}\hat{b}^r_\mathbf{q}v^{s\dagger}(\mathbf{p})v^r(\mathbf{q})e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}})\end{aligned}$$

Now if we perform the position integration first, due to the exponentials, we would get $\delta^{(3)}(\mathbf{p}\pm\mathbf{q})$ from each term and then if we perform the integration over $\mathbf{q}$, these Dirac deltas would get removed, giving $q = \pm p$. Doing this, our integral becomes:

$$\begin{aligned}\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2}\sum_r(\hat{a}^{r\dagger}_\mathbf{p}\hat{a}^r_\mathbf{p}u^{r\dagger}(\mathbf{p})u^r(\mathbf{p}) - \hat{a}^{r\dagger}_\mathbf{p}\hat{b}^r_{-\mathbf{p}}u^{r\dagger}(\mathbf{p})v^r(-\mathbf{p}) \\ +\hat{b}^{r\dagger}_\mathbf{p}\hat{a}^r_{-\mathbf{p}}v^{r\dagger}(\mathbf{p})u^r(-\mathbf{p}) - \hat{b}^{r\dagger}_\mathbf{p}\hat{b}^r_\mathbf{p}v^{r\dagger}(\mathbf{p})v^r(\mathbf{p}))\end{aligned}$$

From Eq. 21 and 22, we obtain a nice, cute equation (well, hidden under the cuteness is a huge mess which we have to deal with),

$$\boxed{\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}E_\mathbf{p}\sum_r\left(\hat{a}^{r\dagger}_\mathbf{p}\hat{a}^r_\mathbf{p} - \hat{b}^{r\dagger}_\mathbf{p}\hat{b}^r_\mathbf{p}\right)} \tag{29}$$

24.5. Commutations?

Till now we have not talked about any commutation relations that is neccessary for quantisation. In general, we assume the field components indexed by $\{\alpha\}$ running from 1 to 4 (since the spinors are basically column vectors) and with our previous intuition, let us guess the commutation relations between the field $\Psi_\alpha(\mathbf{x})$ and the conjugate momentum $i\Psi^\dagger_\alpha(\mathbf{x})$,

$$\begin{aligned} \left[\Psi_\alpha(\mathbf{x}), \Psi_\beta(\mathbf{y})\right] &= \left[\Psi^\dagger_\alpha(\mathbf{x}), \Psi^\dagger_\beta(\mathbf{y})\right] = 0 \\ \left[\Psi_\alpha(\mathbf{x}), \Psi^\dagger_\beta(\mathbf{y})\right] &= \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{y}) \end{aligned} \tag{30}$$

Let us see the commutations between the coefficient operators $\hat{a}$ and $\hat{b}$. We first claim the expressions and show that these indeed conform to the actual field operator commutators.

$$\begin{aligned} \left[\hat{a}^s_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\right] &= (2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) \\ \left[\hat{b}^s_{\mathbf{p}}, \hat{b}^{r\dagger}_{\mathbf{q}}\right] &= (2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) \end{aligned} \tag{31}$$

with all other commutators vanishing. We now check whether using the above relations, we recover the field commutations from Eq. 30 or not. For simplicity, we will suppress the index for the components of the field.

$$\begin{aligned} \left[\Psi(\mathbf{x}), \Psi^\dagger(\mathbf{y})\right] &= \sum_{r,s}\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_{\mathbf{p}}E_{\mathbf{q}}}}\Big(\left[\hat{a}^r_{\mathbf{p}}, \hat{a}^{s\dagger}_{\mathbf{q}}\right]u^r(\mathbf{p})u^{s\dagger}(\mathbf{q})e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} \\ &\qquad + \left[\hat{b}^r_{\mathbf{p}}, \hat{b}^{s\dagger}_{\mathbf{q}}\right]v^r(\mathbf{p})v^{s\dagger}(\mathbf{q})e^{i(\mathbf{q}\cdot\mathbf{y}-\mathbf{p}\cdot\mathbf{x})}\Big) \\ &= \sum_r\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\mathrm{d}^3\mathbf{q}\frac{1}{2\sqrt{E_{\mathbf{p}}E_{\mathbf{q}}}}\delta^{(3)}(\mathbf{p}-\mathbf{q})\Big(u^r(\mathbf{p})u^{r\dagger}(\mathbf{q})e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + v^r(\mathbf{p})v^{r\dagger}(\mathbf{q})e^{i(\mathbf{q}\cdot\mathbf{y}-\mathbf{p}\cdot\mathbf{x})}\Big) \\ &= \sum_r\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\Big(u^r(\mathbf{p})u^{r\dagger}(\mathbf{p})e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + v^r(\mathbf{p})v^{r\dagger}(\mathbf{p})e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\Big) \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\Big((\not{p}+m)\gamma^0e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + (\not{p}-m)\gamma^0e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\Big) \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\left[(\gamma^0E_{\mathbf{p}}+\gamma^ip_i+m)+(\gamma^0E_{\mathbf{p}}-\gamma^ip_i-m)\right]\gamma^0e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} = \delta^{(3)}(\mathbf{x}-\mathbf{y}) \qquad \text{(as expected)} \end{aligned}$$

where in the above steps, we have used Eq. 23 and in the second last step, in the second term, we changed $\mathbf{p}\to-\mathbf{p}$ in the integral, so $p_i\to-p_i$ accordingly. We thus see that assuming these form of creation/annihilation operators yields the correct field quantisation.[1]

Now, note that we see an inocuous $-$ sign in the Hamiltonian (Eq. 29) which imply that the spectrum is unbounded from below. If we create more and more particles with $\hat{b}^\dagger$, we can lower the energy indefinitely which is really, really *unphysical*. We could instead make a reinterpretation as $\hat{b}^\dagger \leftrightarrow \hat{b}$ but then the commutation relations in Eq. 31 gets ruined and the states now get negative norm.

The fundamental flaw was that the Dirac equation describes fermions which follow entirely different statistics than bosons (described by KG field from which we took the inspiration for the commutator).

[1] Technically, the commutation result should give

$$\left[\Psi_\alpha(\mathbf{x}), \Psi^\dagger_\beta(\mathbf{y})\right] = \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{y})$$

if we have used the index for the field components. Hence the commutation relation between the fields is proportional to the identity matrix.

Lecture 25: Quantisation of Dirac Field: What to do!

In the last section, we saw some problems regarding the Hamiltonian which came from assuming a commutation relation for fermionic fields described by the Dirac equation. Fermions follow a different statistics altogether, being constrained by the *exclusion principle*.

25.1. Quantising Fermions

The fermionic wavefunction should be anti-symmetric under exchange of particles. Hence, instead of commutation relation, it is better to consider anti-commutation relation between the fields.

$$\{\Psi_\alpha(\mathbf{x}), \Psi^\dagger_\beta(\mathbf{y})\} = \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{y})$$

and all other anti-commutations vanish. With these anti-commutation relations, we find a suitable relation between the field creation/annihilation operators.

$$\{\hat{a}^r_{\mathbf{p}}, \hat{a}^{s\dagger}_{\mathbf{q}}\} = \delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) \qquad \{\hat{b}^r_{\mathbf{p}}, \hat{b}^{s\dagger}_{\mathbf{q}}\} = \delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) \tag{32}$$

with all other anti-commutations vanishing. Notice that while deriving the Hamiltonian we did not use the commutation relation and thus, the Hamiltonian still remains the same,

$$\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} \sum_r \left(\hat{a}^{r\dagger}_{\mathbf{p}}\hat{a}^r_{\mathbf{p}} - \hat{b}^{r\dagger}_{\mathbf{p}}\hat{b}^r_{\mathbf{p}}\right)$$

But now, let us now make the reinterpretation $\hat{b}^\dagger \leftrightarrow \hat{b}$ and then we use the anti-commutation relation:

$$\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} \sum_r \left(\hat{a}^{r\dagger}_{\mathbf{p}}\hat{a}^r_{\mathbf{p}} + \hat{b}^{r\dagger}_{\mathbf{p}}\hat{b}^r_{\mathbf{p}} - (2\pi)^3\delta^{(3)}(0)\right)$$

We see that the negative sign now has been removed. The ominous infinite term is successfully neglected (or more sophisticatedly removed by *normal ordering*) which we had done for the KG field too. The total Hamiltonian now becomes:

$$\boxed{\hat{H} = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} \sum_r \left(\hat{a}^{r\dagger}_{\mathbf{p}}\hat{a}^r_{\mathbf{p}} + \hat{b}^{r\dagger}_{\mathbf{p}}\hat{b}^r_{\mathbf{p}}\right)}$$

25.2. Constructing the Hilbert space

As with the KG field, we will now construct multi-particle states here also. Define the vacuum state $|\Omega\rangle$ such that,

$$\hat{a}^s_{\mathbf{p}}|\Omega\rangle = 0 \qquad \hat{b}^s_{\mathbf{p}}|\Omega\rangle = 0 \ \forall\ \mathbf{p}, s$$

Then $\hat{H}|\Omega\rangle = 0$ as it should be. Let us now calculate some commutation relations between the Hamiltonian and the creation/annihilation operators which will allow us to create the other states. First note that,

$$\left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}} - \hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}}\right] = \left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}\right]$$

since $\hat{b}$ and $\hat{b}^\dagger$ anti-commute with $\hat{a}$, we have:

$$(\hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}})\hat{a}^s_{\mathbf{p}} = -\hat{b}^{r\dagger}_{\mathbf{q}}\hat{a}^s_{\mathbf{p}}\hat{b}^r_{\mathbf{q}} = \hat{a}^s_{\mathbf{p}}(\hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}}) \implies \left[\hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}}, \hat{a}^s_{\mathbf{p}}\right] = 0$$

and same happens with $\hat{a}^\dagger$. Then, we expand the commutator:

$$\begin{aligned}\left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}\right] = \hat{a}^{s\dagger}_{\mathbf{p}}(\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}) - (\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}})\hat{a}^{s\dagger}_{\mathbf{p}} = -\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^{s\dagger}_{\mathbf{p}}\hat{a}^r_{\mathbf{q}} - (\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}})\hat{a}^{s\dagger}_{\mathbf{p}} &= -\hat{a}^{r\dagger}_{\mathbf{q}}\left(\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) - \hat{a}^r_{\mathbf{q}}\hat{a}^{s\dagger}_{\mathbf{p}}\right) - (\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}})\hat{a}^{s\dagger}_{\mathbf{p}} \\ &= -\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\hat{a}^{r\dagger}_{\mathbf{q}}\end{aligned}$$

From this, we have the commutation with the Hamiltonian as:

$$\left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{H}\right] = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} E_{\mathbf{q}} \sum_r \left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}} - \hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}}\right] = -\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} E_{\mathbf{q}} \delta^{(3)}(\mathbf{p}-\mathbf{q})\hat{a}^{s\dagger}_{\mathbf{q}} = -E_{\mathbf{p}}\hat{a}^{s\dagger}_{\mathbf{p}}$$

With a similar philosophy, we have:

$$\left[\hat{a}^{s}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}} - \hat{b}^{r\dagger}_{\mathbf{q}}\hat{b}^r_{\mathbf{q}}\right] = \left[\hat{a}^{s}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}\right]$$

from following the calculation below,

$$\begin{aligned}\left[\hat{a}^{s}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}\right] = \hat{a}^{s}_{\mathbf{p}}(\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}) - (\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}})\hat{a}^{s}_{\mathbf{p}} = \hat{a}^{s}_{\mathbf{p}}(\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}) + \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^{s}_{\mathbf{p}}\hat{a}^r_{\mathbf{q}} &= \hat{a}^{s}_{\mathbf{p}}(\hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}) + (\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) - \hat{a}^{s}_{\mathbf{p}}\hat{a}^{r\dagger}_{\mathbf{q}})\hat{a}^r_{\mathbf{q}} \\ &= \delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\hat{a}^r_{\mathbf{q}}\end{aligned}$$

From this, we have the commutation with the Hamiltonian as:

$$\left[\hat{a}^{s}_{\mathbf{p}}, \hat{H}\right] = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} E_{\mathbf{q}} \sum_r \left[\hat{a}^{s}_{\mathbf{p}}, \hat{a}^{r\dagger}_{\mathbf{q}}\hat{a}^r_{\mathbf{q}}\right] = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} E_{\mathbf{q}} \delta^{(3)}(\mathbf{p}-\mathbf{q})\hat{a}^{s}_{\mathbf{q}} = E_{\mathbf{p}}\hat{a}^{s}_{\mathbf{p}}$$

Since $\hat{b}$ is identical to $\hat{a}$, it has the same commutation with the Hamiltonian. We finally have:

$$\triangleright \left[\hat{a}^{s}_{\mathbf{p}}, \hat{H}\right] = E_{\mathbf{p}}\hat{a}^{s}_{\mathbf{p}} \quad \triangleright \left[\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{H}\right] = -E_{\mathbf{p}}\hat{a}^{s\dagger}_{\mathbf{p}} \quad \triangleright \left[\hat{b}^{s}_{\mathbf{p}}, \hat{H}\right] = E_{\mathbf{p}}\hat{b}^{s}_{\mathbf{p}} \quad \triangleright \left[\hat{b}^{s\dagger}_{\mathbf{p}}, \hat{H}\right] = -E_{\mathbf{p}}\hat{b}^{s\dagger}_{\mathbf{p}}$$

Now let us define a state given by:

$$|\mathbf{p}, s\rangle = \sqrt{2E_{\mathbf{p}}}\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle$$

Let us see what happens when we apply the Hamiltonian on it. We use the commutation relations found above:

$$E_{\mathbf{p}}(\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle) = \left[\hat{H}, \hat{a}^{s\dagger}_{\mathbf{p}}\right]|\Omega\rangle = \hat{H}(\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle) - 0 \implies \hat{H}(\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle) = E_{\mathbf{p}}(\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle)$$

Thus we see that $|\mathbf{p}, s\rangle$ has energy $E_{\mathbf{p}}$. By constructing the momentum operator, we can also verify that this state indeed has momentum $\mathbf{p}$ and thus, we can safely interpret this as a *one-particle* state (*fermions*). Similar thing occurs for anti-particle, which is created by $\hat{b}^{\dagger}$.

$$|\tilde{p}, s\rangle = \sqrt{2E_{\mathbf{p}}}\hat{b}^{s\dagger}_{\mathbf{p}}|\Omega\rangle$$

where the tilde denotes that it is an anti-particle (*anti-fermions*). Anti-particles are also having the same energy $E_{\mathbf{p}}$.

25.3. Pauli's Exclusion Principle

A characteristic property of fermions is the *exclusion principle* which roughly states that two fermions with the same quantum numbers cannot occupy the same place. Let us check if this property is satisfied by our theory or not. For that, note:

$$\{\hat{a}^{s\dagger}_{\mathbf{p}}, \hat{a}^{s\dagger}_{\mathbf{p}}\} = 0 \implies (\hat{a}^{s\dagger}_{\mathbf{p}})^2 \equiv 0$$

Now, if we create a two fermion state as:

$$|\mathbf{p}, \mathbf{p}; s, s\rangle = \sqrt{2E_{\mathbf{p}}}\sqrt{2E_{\mathbf{p}}}\hat{a}^{s\dagger}_{\mathbf{p}}\hat{a}^{s\dagger}_{\mathbf{p}}|\Omega\rangle \sim (\hat{a}^{s\dagger}_{\mathbf{p}})^2|\Omega\rangle = 0$$

which nicely exhibits Pauli's exclusion principle and thus, this is automatically incorporated into the theory through the anti-commutation relations.

Lecture 26: Gauging the symmetry!

Here we will consider some non-local symmetries where the parameter itself depends on spacetime. Remember in sec. 9.2.1, we had considered a global U(1) symmetry for the complex scalar field, by the transformation $\phi \to e^{i\alpha}\phi$. There, α was a global parameter, which did not depend on spacetime. Here we will consider the generalisation when α can indeed depend on space-time, giving rise to *local symmetry.* Consider the same Lagrangian as before,

$$\mathcal{L} = (\partial^\mu \varphi^*)(\partial_\mu \varphi) - m^2 \varphi^* \varphi$$

Suppose $\varphi(x) \to e^{iq\mathbb{e}\alpha(x)}\varphi(x)$ such that $\alpha(x)$ is now spacetime dependent and the parameter of transformation is $q\mathbb{e}$, with q being called the *coupling constant.* The second term is invariant under this transformation and the problem arises due to the first (kinetic) term which does not remain invariant under the transformation,

$$\partial_\mu \varphi(x) \longrightarrow \partial_\mu \left(e^{iq\mathbb{e}\alpha(x)}\varphi\right) = (\partial_\mu \varphi(x) + iq\mathbb{e}\ \partial_\mu \alpha(x)\varphi(x))e^{iq\mathbb{e}\alpha(x)}$$

To resolve the problem, we will define a new kind of derivative, called the *gauge covariant derivative*,

$$\mathfrak{D}_\mu \varphi := \partial_\mu \varphi + \Gamma_\mu \varphi \qquad \Gamma_\mu \equiv -iq\mathbb{e}A_\mu(x)$$

where $A_\mu(x)$ is called the *gauge field* and its transformation is defined as,

$$A_\mu(x) \longrightarrow A_\mu(x) + \partial_\mu \alpha(x)$$

With the new definition of the derivative, we can see

$$\begin{aligned}
\mathfrak{D}_\mu \varphi(x) \longrightarrow & [\partial_\mu - iq\mathbb{e}(A_\mu(x) + \partial_\mu \alpha(x))]e^{iq\mathbb{e}\alpha(x)}\varphi(x) \\
= & [(\partial_\mu \varphi(x) + \cancel{iq\mathbb{e}\ \partial_\mu \alpha(x)\varphi(x)}) - \cancel{iq\mathbb{e}\partial_\mu \alpha(x)\varphi x} - iq\mathbb{e}A_\mu(x)\varphi(x)]e^{iq\mathbb{e}\alpha(x)} \\
= & e^{iq\mathbb{e}\alpha(x)}[\partial_\mu \varphi(x) - iq\mathbb{e}A_\mu(x)]\varphi \\
= & e^{iq\mathbb{e}\alpha(x)}\mathfrak{D}_\mu \varphi(x) \implies (\mathfrak{D}_\mu \varphi(x))^* \longrightarrow e^{-iq\mathbb{e}\alpha(x)}(\mathfrak{D}_\mu \varphi(x))^*
\end{aligned}$$

Hence the gauge covariant derivative transforms as we intended, making $\mathcal{L}$ invariant under the local U(1) gauge transformation. This kind of treatment is called *gauging the symmetry*, more specifically, we are considering an *abelian* gauge theory, since the symmetry transformation commutes with the field. Henceforth we replace all normal derivatives in the Lagrangian with the covariant derivative, doing which we obtain

$$\begin{aligned}
\mathcal{L} &= (\mathfrak{D}^\mu \varphi^*)(\mathfrak{D}_\mu \varphi) - m^2 \varphi^* \varphi \\
&= (\partial^\mu + iq\mathbb{e}A^\mu)\varphi^*(\partial_\mu - iq\mathbb{e}A_\mu)\varphi - m^2 \varphi^* \varphi \\
&= (\partial^\mu \varphi^* + iq\mathbb{e}A^\mu \varphi^*)(\partial_\mu \varphi - iq\mathbb{e}A_\mu \varphi) - m^2 \varphi^* \varphi \\
&= \left[(\partial^\mu \varphi^*)(\partial_\mu \varphi) - m^2 \varphi^* \varphi\right] - iq\mathbb{e}A_\mu \varphi \partial^\mu \varphi^* + iq\mathbb{e}A^\mu \varphi^* \partial_\mu \varphi + g^2\mathbb{e}^2 \varphi^* \varphi A^\mu A_\mu \\
&= \left[(\partial^\mu \varphi^*)(\partial_\mu \varphi) - m^2 \varphi^* \varphi\right] - q\mathbb{e}A^\mu i(\varphi \partial_\mu \varphi^* - \varphi^* \partial_\mu \varphi) + g^2\mathbb{e}^2 \varphi^* \varphi A^\mu A_\mu
\end{aligned}$$

The blue term is the original scalar field and note from Eq. 2, the red term is actually the conserved Noether current J_μ. Thus, we see that the gauge field now couples to the Noether through the red term and green terms is the four-point interaction term, often called *seagull term* since the Feynman diagram (to be discuss later) looks like seagulls flying with their wings.

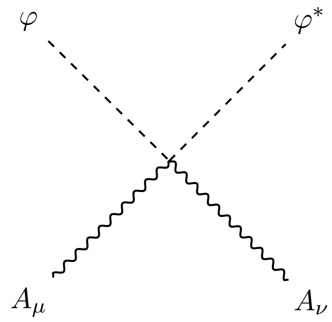


Figure 7: Feynman diagram showing the 'seagull' term. The dashed lines represent the scalar field and the wiggly lines represent the guage field.

26.1. Field strength

To make the gauge field a true dynamical variable we need to add a term to the Lagrangian involving its derivatives, since the EoM are found from the derivatives. We thus define the *field strength*,

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$$

The field strength is anti-symmetric in its indices. Considering that normal partial derivatives commute, under the transformation, the field strength remains invariant,

$$F_{\mu\nu} \longrightarrow \partial_\mu(A_\nu + \partial_\nu \alpha(x)) - \partial_\nu(A_\mu + \partial_\mu \alpha(x)) = F_{\mu\nu}$$

The simplest gauge-invariant Lagrangian we can make out of the field strength is $\mathcal{L}_s = -1/4 F_{\mu\nu}F^{\mu\nu}$ and thus, we get the total Lagrangian as,

$$\boxed{\mathcal{L} = (\mathfrak{D}^\mu \varphi^*)(\mathfrak{D}_\mu \varphi) - m^2 \varphi^* \varphi - \frac{1}{4} F_{\mu\nu}F^{\mu\nu}} \tag{33}$$

26.2. Gauging the Dirac field

We will now couple the gauge field to the Dirac field. The Dirac Lagrangian looks very similar to the complex scalar field Lagrangian, so here also we will replace the normal derivative with the covariant derivative and this is indeed invariant under the local gauge transformation as $\overline{\Psi}(x) = \Psi^\dagger(x)\gamma^0$ here has the same effect as $\varphi^*(x)$ on the transformation.

$$\boxed{\mathcal{L}_\text{D} = i\overline{\Psi}\gamma^\mu \mathfrak{D}_\mu \Psi - m\overline{\Psi}\Psi - \frac{1}{4} F_{\mu\nu}F^{\mu\nu}}$$

Expanding the covariant derivative as before, we get

$$\begin{aligned}
\mathcal{L}_\text{D} &= i\overline{\Psi}\gamma^\mu(\partial_\mu - iq\mathbb{e}A_\mu(x))\Psi - m\overline{\Psi}\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
&= [i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m\overline{\Psi}\Psi] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + q\mathbb{e}\overline{\Psi}\gamma^\mu\Psi A_\mu(x)
\end{aligned}$$

The blue term, as before denotes the free Dirac field. From Eq. 28, we see that the red term denotes the conserved Noether current which gets coupled to the gauge field, like the previous case.

If we interpret the field strenth as the electromagnetic field tensor and A_μ as the four-potential, the above modified Lagrangian essentially describes the coupling of a photon to an electron, which forms the basis of the quantum electrodynamics (QED).

26.3. Equations of Motion

In both the cases above, we saw that expanding the kinetic term led to the coupling between the gauge field and the current. Thus, if we consider a generic Lagrangian,

$$\mathcal{L}_{\text{g}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J^{\mu}A_{\mu}$$

Since we are focussing on the dynamics of the gauge field, we need to find the EoM using the gauge field, that is,

$$\partial_{\mu}\left(\frac{\partial\mathcal{L}_{\text{g}}}{\partial(\partial_{\mu}A_{\nu})}\right) - \underbrace{\frac{\partial\mathcal{L}_{\text{g}}}{\partial A_{\nu}}}_{-J^{\nu}} = 0$$

Let us calculate this step-by-step by first noting that

$$\begin{aligned}
F_{\mu\nu}F^{\mu\nu} &= (\partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu})(\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}) \\
&= \partial_{\mu}A_{\nu}\partial^{\mu}A^{\nu} - \partial_{\nu}A_{\mu}\partial^{\mu}A^{\nu} - \partial_{\mu}A_{\nu}\partial^{\nu}A^{\mu} + \partial_{\nu}A_{\mu}\partial^{\nu}A^{\mu} \\
&= 2[\partial_{\mu}A_{\nu}\partial^{\mu}A^{\nu} - \partial_{\nu}A_{\mu}\partial^{\mu}A^{\nu}] \\
&= 2[\partial_{\mu}A_{\nu}\partial^{\mu}A^{\nu} - \partial_{\mu}A_{\nu}\partial^{\nu}A^{\mu}] \\
&= 2\ \partial_{\mu}A_{\nu}F^{\mu\nu}
\end{aligned} \tag{34}$$

Also note that,

$$\frac{\partial F^{\rho\sigma}}{\partial(\partial_{\mu}A_{\nu})} = \frac{\partial(\partial^{\rho}A^{\sigma})}{\partial(\partial_{\mu}A_{\nu})} - \frac{\partial(\partial^{\sigma}A^{\rho})}{\partial(\partial_{\mu}A_{\nu})} = (\eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu})$$

Using this, we get the equation of motion as:

$$\begin{aligned}
0 &= -\frac{1}{4}\partial_{\mu}\left(\frac{\partial(F_{\rho\sigma}F^{\rho\sigma})}{\partial(\partial_{\mu}A_{\nu})}\right) + J^{\nu} \\
&= -\frac{1}{4} \times 2 \times \partial_{\mu}\left[\delta^{\mu}{}_{\rho}\delta^{\nu}{}_{\sigma}F^{\rho\sigma} + \partial_{\rho}A_{\sigma}\frac{\partial F^{\rho\sigma}}{\partial(\partial_{\mu}\partial_{\nu})}\right] + J^{\nu} \\
&= -\frac{1}{2} \times \partial_{\mu}[F^{\mu\nu} + \partial_{\rho}A_{\sigma}(\eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu})] + J^{\nu} \\
&= -\frac{1}{2} \times \partial_{\mu}[F^{\mu\nu} + \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}] + J^{\nu} \\
&= -\ \partial_{\mu}F^{\mu\nu} + J^{\nu}
\end{aligned}$$

We get the equation of motion as very similar to the Maxwell's equations,

$$\boxed{\partial_{\mu}F^{\mu\nu} = J^{\nu}} \tag{35}$$

Let us now find the conjugate momenta for the gauge field

$$\Pi^{\mu} = \frac{\partial\mathcal{L}_{\text{g}}}{\partial(\partial_0 A_{\mu})} = -F^{0\mu}$$

- ▷ $\mu = 0$ (temporal component) $\longrightarrow \Pi^0 = -F^{00} = 0$ which is a *primary constraint* since this does not depend on the EoM. This results purely from the Lagrangian definition.
- ▷ $\mu \neq 0$ (spatial components) $\longrightarrow \Pi^i = -F^{0i} = E^i$ from sec B.

where E^i denotes the component of the electric field. Thus, we have the conjugate momenta as $\Pi^{\mu} = (0, \mathbf{E})$.

26.4. Solution to the Free Gauge field

As seen above, the gauge field satisfies the wave-equation for the Feynman gauge and we can expand the gauge field using plane waves

$$A^\mu(x) = \epsilon^\mu(p)e^{-ipx}$$

where $\epsilon^\mu(x)$ are called *polarisation vectors*. Substituting this in the EoM $\partial_\mu J^{\mu\nu} = 0$, we get

$$\begin{aligned}0 &= \partial_\mu\left[\partial^\mu(\epsilon^\nu(p)e^{-ipx}) - \partial^\nu(\epsilon^\mu(p)e^{-ipx})\right]\\ &= \partial_\mu\left[\epsilon^\nu(p)(-ip^\mu)e^{-ipx} - (-ip^\nu)\epsilon^\mu(p)e^{-ipx}\right]\\ &= (-p^2\epsilon^\nu(p) + p_\mu p^\nu\epsilon^\mu(p))e^{-ipx}\\ &= (-p^2\eta^{\nu\mu}\epsilon_\mu(p) + p^\mu p^\nu\epsilon_\mu(p))e^{-ipx}\\ &= -(p^2\eta^{\nu\mu} - p^\mu p^\nu)\epsilon_\mu(p)e^{-ipx}\end{aligned}$$

Let us define the 4×4 matrix $\mathcal{M}^{\mu\nu} = p^2\eta^{\mu\nu} - p^\mu p^\nu$ from which we get $\mathcal{M}^{\mu\nu}\epsilon_\mu(p) = 0$. Let us define the quantities

$$P^{\mu\nu} = \eta^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \qquad Q^{\mu\nu} = \frac{p^\mu p^\nu}{p^2}$$

Now, we check some properties of these operators

- ▷ Idempotence: (kinda?)

$$\begin{aligned}P^\mu{}_\beta P^{\beta\nu} &= P^{\mu\alpha}g_{\alpha\beta}P^{\beta\nu}\\ &= \left(\eta^{\mu\alpha} - \frac{p^\mu p^\alpha}{p^2}\right)\eta_{\alpha\beta}\left(\eta^{\beta\nu} - \frac{p^\beta p^\nu}{p^2}\right)\\ &= \eta^{\mu\nu} - \delta^\mu{}_\beta\frac{p^\beta p^\nu}{p^2} - \delta^\nu{}_\alpha\frac{p^\mu p^\alpha}{p^2} + \frac{p^\mu p^\alpha p^\beta p^\nu}{p^4}\eta_{\alpha\beta}\\ &= \eta^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} - \frac{p^\mu p^\nu}{p^2} + \frac{p^\mu p^\nu}{p^2} = P^{\mu\nu}\\ Q^\mu{}_\beta Q^{\beta\nu} &= \frac{p^\mu p_\beta}{p^2}\frac{p^\beta p_\nu}{p^2} = \frac{p^\mu p^\nu}{p^2} = Q^{\mu\nu}\end{aligned}$$

- ▷ Orthogonality: $Q^\mu{}_\beta P^{\beta\nu} = \frac{p^\mu p_\beta}{p^2}\left(\eta^{\beta\nu} - \frac{p^\beta p^\nu}{p^2}\right) = \frac{p^\mu p^\nu}{p^2} - \frac{p^\mu p^\nu}{p^2} = 0$
- ▷ Completition: $P^{\mu\nu} + Q^{\mu\nu} = \eta^{\mu\nu}$

Thus we can vaguely identify these quantities as some kind of projection operators. Let $P^{\mu\nu}$ act on some polarization vector, say ϵ_μ,

$$P^{\mu\nu}\epsilon_\nu = \epsilon^\mu - \frac{p^\mu(p^\nu\epsilon_\nu)}{p^2} \implies p_\mu(P^{\mu\nu}\epsilon_\nu) = p_\mu\epsilon^\mu - p_\nu\epsilon^\nu = 0$$

Thus $P^{\mu\nu}$ projects the polarization vector onto the subspace orthogonal to the momentum p^μ.

Lecture 27: Quantising the gauge

In the previous section, we had an introduction to local gauge transformations and now, we will proceed to understand some subtleties regarding the gauge fields and ultimately find ways to quantise them. To start with, let us see gauge transformations a bit mathematically.

Consider the spacetime (which acts as a *base* for our theory) as $\mathcal{M}$ and the group U(1) (since we are considering local $U(1)$ transformations). Associated to the this group is a connection A_μ, also called *gauge field* and let $\mathcal{A} = \{A_\mu\}$ be the collection of all such gauge fields.

Now, consider the group $\mathcal{G}$ of guage transformations, defined as

$$\mathcal{G} = \{g|\ g : \mathcal{M} \to G\}$$

An element $g \in \mathcal{G}$ acts on the connection as

$$A_\mu \mapsto gA_\mu g^{-1} - g\partial_\mu g^{-1}$$

Given a gauge field A_μ we can define set of all gauge transformations on the field, which is called the set of *gauge orbits* $\mathcal{O}_{\mathrm{A}}$

$$\mathcal{O}_{\mathrm{A}} := \{gA_\mu|\ g \in \mathcal{G}\}$$

Let us see the example for U(1) symmetry. The set of gauge transformations $\mathcal{G}$ is such that for $g \in \mathcal{G}$,

$$g_\alpha(x) = e^{iq\mathbb{e}\alpha(x)}\ \forall\ x \in \mathcal{M}$$

Then we have $g_\alpha A_\mu g_\alpha^{-1} = A_\mu$ since g_α commutes with the gauge field. Now, for the derivative we have,

$$g_\alpha \partial_\mu g_\alpha^{-1} = e^{iq\mathbb{e}\alpha(x)}\partial_\mu(e^{-iq\mathbb{e}\alpha(x)}) = -iq\mathbb{e}\partial_\mu\alpha$$

which is exactly the transformation that we had seen earlier (we incorporated $iq\mathbb{e}$ in the covariant derivative definition explicitly, so the factor here makes sense).

Each element in the orbit is called *gauge-equivalent* since all these are related by some gauge transformation. The full space $\mathcal{A}$ is thus redundant since there are so many A_μ's which are connected by gauge transformations. Hence, choose some representative point from each gauge orbit and describe our physical observables with respect to these representative points, which is called *gauge-fixing*.

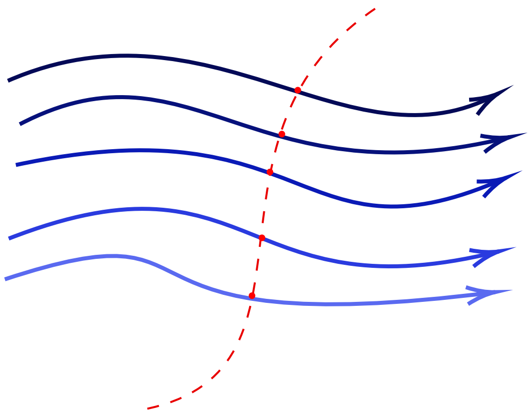


Figure 8: Fixing the gauge. Each continuous lines denote the gauge orbits and the red points are the representative points denoting the gauge fixing condition

The general theory of gauge contains rigorous concepts of differential geometry, like principle fibre bundles and sections defined over them, which is *extremely interesting but of not much use for physical purposes*[1]. The basic essence of gauge-fixing is to reduce the degrees of freedom.

Two most popular choices for gauge-fixing are the *Lorenz gauge* where the representative A_μ satisfies $\partial_\mu A^\mu = 0$ and the Coulomb gauge, where the representative A_μ satsifies $\nabla \cdot \mathbf{A} = 0$.

Note that the Lorenz gauge does not uniquely determine a representative from the gauge orbit since if A_μ satisfies the condition, then for any λ satisfying $\partial_\mu\partial^\mu\lambda = 0$ (that is, the wave-equation), $A_\mu + \partial_\mu\lambda$ also satisfies the gauge-fixing condition. Hence in the Lorenz gauge, there is always some residual gauge freedom left.

[1] just like me

27.1. Fixing the Gauge

We will try to impose the gauge fixing condition by modifying the free electromagnetic field Lagrangian with an additional term

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

where ξ is a generic parameter, which weighs gauge equivalent configurations differently. We will later impose the condition that $\partial_\mu A^\mu$ vanishes between physical states. Let us expand the Lagrangian a bit. The first term as seen from Eq. 34 gives $-\frac{1}{2}\partial_\mu A_\nu F^{\mu\nu}$. Let us focus on the second term carefully. We can write,

$$\begin{aligned}(\partial_\mu A^\mu)(\partial_\nu A^\nu) &= \partial_\mu[A^\mu(\partial_\nu A^\nu)] - A^\mu\partial_\mu(\partial_\nu A^\nu) \\ &= \partial_\mu[A^\mu(\partial_\nu A^\nu)] - A^\mu\partial_\nu(\partial_\mu A^\nu) \\ &= \partial_\mu[A^\mu(\partial_\nu A^\nu)] - [\partial_\nu(A^\mu(\partial_\mu A^\nu)) - (\partial_\mu A^\nu)(\partial_\nu A^\mu)] \\ &= (\partial_\mu A^\nu)(\partial_\nu A^\mu) + \partial_\mu[A^\mu(\partial_\nu A^\nu) - A^\nu\partial_\nu A^\mu]\end{aligned}$$

The term in the parenthesis is a total derivative and in the action, this will vanish at the boundaries, without affecting the EoM. Hence, we can neglect this term and the Lagrangian can be written as,

$$\begin{aligned}\mathcal{L}_{\text{EM}} &= -\frac{1}{2}\partial_\mu A_\nu F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\nu)(\partial_\nu A^\mu) \\ &= -\frac{1}{2}\left[(\partial_\mu A_\nu)(\partial^\mu A^\nu) - (\partial_\mu A_\nu)(\partial^\nu A^\mu) + \frac{1}{\xi}(\partial_\mu A^\nu)(\partial_\nu A^\mu)\right] \\ &= -\frac{1}{2}\left[(\partial_\mu A_\nu)(\partial^\mu A^\nu) - \left(1-\frac{1}{\xi}\right)(\partial_\mu A_\nu)(\partial^\nu A^\mu)\right]\end{aligned}$$

From this, we can find the equation of motion.

$$0 = \partial_\sigma\left(\frac{\partial\mathcal{L}_{\text{EM}}}{\partial(\partial_\sigma A_\rho)}\right) - \frac{\partial\mathcal{L}_{\text{EM}}}{\partial A_\sigma} \implies \boxed{\partial_\sigma\partial^\sigma A^\rho - \left(1-\frac{1}{\xi}\right)\partial^\rho\partial_\sigma A^\sigma = 0} \tag{36}$$

In terms of $F^{\mu\nu}$ we get the EoM as

$$\partial_\mu F^{\mu\nu} + \frac{1}{\xi}\partial^\nu(\partial_\alpha A^\alpha) = 0$$

For $\xi = 1$, $\partial_\sigma\partial^\sigma A^\rho = 0$ which admits a plane-wave solution for the gauge field. This choice is termed as *Feynman gauge* and calculations are easy with this choice. Another choice is to take $\xi \to 0$ which is called the *Landau gauge*.

Now let us find the canonical momentum for the Lagrangian. We will first find it from the proposed Lagrangian, without the boudary modification. The momentum from the first term, $\Pi^\mu = F^{-0\mu}$, was already seen before. Noting that $\partial_\mu A^\mu = (\partial_0 A^0 + \partial_i A^i)$, from the second term, we have

$$(\Pi^\rho)_2 = -\frac{1}{2\xi}[2\times\delta^\rho_0\partial_\mu A^\mu] = -\frac{1}{\xi}\delta^\rho_0\partial_\mu A^\mu \longrightarrow \left(\Pi^0\right)_2 = -\frac{1}{\xi}\partial_\mu A^\mu \quad \left(\Pi^i\right)_2 = 0$$

Then, the conjugate momentum for this modified Lagrangian becomes:

$$\Pi^0 = -\frac{1}{\xi}\partial_\mu A^\mu \qquad \Pi^i = -F^{0i}$$

Now if we consider the Lagrangian where we removed the total derivative term, we get an alternate conjugate momentum

$$\Pi^\rho := \frac{\partial\mathcal{L}_{\text{EM}}}{\partial(\partial_0 A_\rho)} = -\partial^0 A^\rho + \left(1-\frac{1}{\xi}\right)\partial^\rho A^0$$

In this case, we see that $\Pi^0 \neq 0$ and thus, the *primary constraint* is gone.

27.2. Quantisation in Feynman Gauge

We impose the usual commutation relation between the fields and the conjugate momentum when elevating them to operators

$$[A^\mu(x), A^\nu(y)] = 0 \qquad [\Pi^\mu(x), \Pi^\nu(y)] = 0 \qquad [A^\mu(x), \Pi^\nu(y)] = i\eta^{\mu\nu}\delta^{(3)}(\mathbf{x}-\mathbf{y})$$

The metric ţ is forced upon us from the condition of Lorentz covariance, since the left-hand side is a tensor. And we can expand these in terms of the creation and annihilation operators, and the polarisation vectors. Since we are considering photons where $m = 0$ we have $p^2 = 0 \implies E_{\mathbf{p}} = |\mathbf{p}|$. Since the polarisation vectors are 4-vectors we write the expansion as a sum of the four components.

$$A_\mu(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2|\mathbf{p}|}}\sum_{\lambda=0}^{3}(\epsilon_\mu)_\lambda \hat{a}^\lambda_{\mathbf{p}} e^{-ipx} + (\epsilon^*_\mu)_\lambda \hat{a}^{\dagger\lambda}_{\mathbf{p}} e^{ipx} \tag{37}$$

We will have four independent solutions for $\epsilon^\mu(p)$ labelled by $\lambda = 0, 1, 2, 3$. We choose a normalisation for the polarisation vectors.

▷ Normalisation: $(\epsilon^\mu)_\lambda(\epsilon^\mu)^*_\lambda \eta_{\mu\nu} = \eta_{\lambda\lambda'}$ For $\mu = 0$ the norm of the polarisation vector is positive while it is negative for $\lambda = 1, 2, 3$.

In a frame where $p^\mu = (p, 0, 0, p)$ we can choose a basis such that $(\epsilon^\mu)_\lambda = \delta^\mu{}_\lambda$ and the expression for any generic frame can be found by an appropriate Lorentz boost. Note that $(\epsilon^\mu)_1$ and $(\epsilon^\mu)_2$ are such that $p_\mu\epsilon^\mu = 0$ that is, these are *transverse* while for $\lambda = 0, 3$ we see that $p_\mu\epsilon^\mu \neq 0$.

Using the canonical momentum expression found out (from the Lagrangian after removing the boundary term) we have:

$$\Pi^\mu(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}(+i\sqrt{\frac{|\mathbf{p}|}{2}})\sum_{\lambda=0}^{3}(\epsilon_\mu)_\lambda \hat{a}^\lambda_{\mathbf{p}} e^{-ipx} - (\epsilon^*_\mu)_\lambda \hat{a}^{\dagger\lambda}_{\mathbf{p}} e^{ipx} \tag{38}$$

We now impose the commutation relations for the creation and annihilation operator

$$\left[\hat{a}^\lambda_{\mathbf{p}}, \hat{a}^{\dagger\lambda'}_{\mathbf{p}}\right] = -\eta^{\lambda\lambda'}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$$

We can verify that indeed, this choice gives us the correct commutation for the field operators.

27.2.1. Finding the Hamiltonian

Again, this is going to be an arduous calculation which feels like a curse (to me). However, I am bored and have nothing better to do right now (this is too frightening that the most exciting thing for me to do right now is to do tedious algebra), so let's do it!

▷ In the Feynman guage we have the Lagrangian density and the momentum as

$$\mathcal{L}_{\text{EM}} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) \qquad \Pi^\mu = -\partial_0 A^\mu$$

▷ The Hamiltonian density is defined as

$$\begin{aligned}
\mathcal{H} = \Pi^\mu\partial_0 A_\mu - \mathcal{L} &= -\Pi^\mu\Pi_\mu + \frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) \\
&= -\Pi^\mu\Pi_\mu + \frac{1}{2}\left[(\partial_0 A_\nu)\left(\partial^0 A^\nu\right) + (\partial_i A_\nu)\left(\partial^i A^\nu\right)\right] \\
&= -\Pi^\mu\Pi_\mu + \frac{1}{2}\Pi^\mu\Pi_\mu - \frac{1}{2}(\partial_i A_\nu)\left(\partial^i A^\nu\right) \\
&= -\frac{1}{2}\Pi^\mu\Pi_\mu + \frac{1}{2}(\partial_i A_\nu)\left(\partial^i A^\nu\right)
\end{aligned}$$

▷ The total Hamiltonian is written as the space integral of the Hamiltonian density.

$$H = \int \mathrm{d}^3x \left[-\frac{1}{2}\Pi^\mu\Pi_\mu + \frac{1}{2}(\partial_i A_\nu)\left(\partial^i A^\nu\right)\right]$$

Then we use the field and canonical momentum expression as in Eq. 37 and Eq. 38. Let us go term by term:
(DO THE INTEGRALS WHENEVER TIME IS THERE!)
We find that the total Hamiltonian is

$$\boxed{H = -\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}|\mathbf{p}| \sum_{\lambda=0}^{3} \eta_{\lambda\lambda}\hat{a}_{\mathbf{p}}^{\dagger\lambda}\hat{a}_{\mathbf{p}}^{\lambda}} \tag{39}$$

The commutation relation between the Hamiltonian and the creation/annihilation operators are thus:

$$\begin{aligned}
\left[H, \hat{a}_{\mathbf{q}}^{\dagger\kappa}\right] &= -\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}|\mathbf{p}| \sum_{\lambda=0}^{3} \eta_{\lambda\lambda}\left[\hat{a}_{\mathbf{p}}^{\dagger\lambda}\hat{a}_{\mathbf{p}}^{\lambda}, \hat{a}_{\mathbf{q}}^{\dagger\kappa}\right] \\
&= -\int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}|\mathbf{p}| \sum_{\lambda=0}^{3} \eta_{\lambda\lambda}\hat{a}_{\mathbf{p}}^{\dagger\lambda}\left[\hat{a}_{\mathbf{p}}^{\lambda}, \hat{a}_{\mathbf{q}}^{\dagger\kappa}\right] \\
&= \int d^3\mathbf{p}\ |\mathbf{p}| \sum_{\lambda=0}^{3} \eta_{\lambda\lambda}\eta^{\lambda\kappa}\hat{a}_{\mathbf{p}}^{\dagger\lambda}\delta^{(3)}(\mathbf{p}-\mathbf{q}) \\
&= |\mathbf{q}|\hat{a}_{\mathbf{p}}^{\dagger\kappa}
\end{aligned}$$

Similarly we also have for the annihilation operator, $\left[H, \hat{a}_{\mathbf{q}}^{\kappa}\right] = -|\mathbf{q}|\hat{a}_{\mathbf{p}}^{\dagger\kappa}$. We could now use these to define the structure of the Hilbert space of the photons.

27.2.2. Fock Space of Photons

We define the vacuum state $|\Omega\rangle$ to be the one killed by the annihilation operator

$$\hat{a}_{\mathbf{p}}^{\lambda}|\Omega\rangle = 0\ \forall\ \mathbf{p}, \lambda$$

We can define the one-particle state with momentum p and polarisation λ by acting the creation operator on the vacuum

$$|\mathbf{p}, \lambda\rangle \equiv \sqrt{2|\mathbf{p}|}\hat{a}_{\mathbf{p}}^{\dagger\lambda}|\Omega\rangle$$

Using this, we can try to calculate the norm of the states with same polarisation.

$$\langle\mathbf{p}, \lambda|\mathbf{q}, \lambda\rangle = 2\sqrt{|\mathbf{p}||\mathbf{q}|}\langle\Omega|\hat{a}_{\mathbf{p}}^{\lambda}\hat{a}_{\mathbf{q}}^{\dagger\lambda}|\Omega\rangle = -2|\mathbf{p}|(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\eta^{\lambda\lambda}$$

using the commutation relation. Tong refers to this as a “wtf” situation and indeed it is, since the state created by the oscillator with $\lambda = 0$ has a negative norm (for $\lambda \neq 0$ it's fine, since $\eta^{\lambda\lambda}$ is itself negative). Since the scalar products in quantum mechanics are interpreted as probabilities, a Fock space with a scalar product which is not positive denite has no probabilistic interpretation.

27.2.3. Gupta-Bleuer Condition

We can resolve this issue by demanding that *physical states* should have positive norms and thus the space-like polarisation ($\lambda = 0$) canot be a physical state. This leads us to the Gupta-Bleuer condition. We consider a subspace $\mathfrak{F}_{\mathrm{p}}$ of the entire Fock space, denoting the physical states and demanding that for states $|\mathrm{phys.}\rangle$ and $|\mathrm{phys.}'\rangle$ belonging to this subspace should satisfy

$$\langle\mathrm{phys.}'|\partial_\mu A^\mu|\mathrm{phys.}\rangle = 0$$

This can be thought of as the use of the Lorenz gauge. Instead of the Lagrangian, the Lorenz gauge is recovered as an operator equation of the physical states.

We first note that the operator $\partial_\mu A^\mu$ can be decomposed into positive and negative frequency parts, that is,

$$(\partial_\mu A^\mu) = (\partial_\mu A^\mu)^+ + (\partial_\mu A^\mu)^-$$

where the corresponding operators are defined as

$$(\partial_\mu A^\mu)^+ = -i \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{2|\mathbf{p}|} \sum_{\lambda=0}^{3} p_\mu (\epsilon^\mu)_\lambda \hat{a}^\lambda_{\mathbf{p}} e^{-ipx}$$

$$(\partial_\mu A^\mu)^- = i \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{2|\mathbf{p}|} \sum_{\lambda=0}^{3} p_\mu (\epsilon^{\mu *})_\lambda \hat{a}^{\dagger\lambda}_{\mathbf{p}} e^{ipx}$$

Since these are Hermitian conjugate of each other, defining the Gupta-Bleuer condition as

$$(\partial_\mu A^\mu)^+ |\text{phys.}\rangle = 0$$

is sufficient since the other condition $(\partial_\mu A^\mu)^- |\text{phys.}\rangle = 0$ is automatically satisfied for any physical state.

The above expression is taken as a definition of the physical state. Note that this definition preserves the linearity of the Hilbert space as using this, the linear combination of any physical state is also physical.

27.2.4. A Few Subtleties...

Let us see what happens to the one-particle state. Consider a general superposition of the one-particle state

$$|\psi\rangle = \sum_\lambda c_\lambda |\mathbf{p}, \lambda\rangle = \sum_\lambda c_\lambda \sqrt{2|\mathbf{p}|} \hat{a}^{\dagger\lambda}_{\mathbf{p}} |\Omega\rangle$$

Choosing a momentum alogn the z-axis $p^\mu = (p, 0, 0, p)$ (and so, $p_\mu = (p, 0, 0, -p)$), applying the Gupta-Bleuer condition and using the commutation relation we will obtain

$$(\partial_\mu A^\mu)^+ |\psi\rangle \propto \sum c_\lambda p_\mu (\epsilon^\mu)^\lambda |\Omega\rangle = 0 \implies c_0 - c_3 = 0$$

Thus c_1 and c_2 can be arbitrary and hence the transverse polarisation states always represent the true physical states. However note that individually $|p, 0\rangle$ and $|p, 3\rangle$ are not physical (since it does not satisfy the Gupta-Bleuer condition), however the combination

$$|\phi\rangle = (\hat{a}^{\dagger 0}_{\mathbf{p}} - \hat{a}^{\dagger 3}_{\mathbf{p}}) |\Omega\rangle$$

is a physical state and hence the most general one-particle state of the physical subspace with momentum $\mathbf{p}$ can be written as,

$$|\psi\rangle = |\psi_{\mathrm{T}}\rangle + \mathcal{C} |\phi\rangle$$

where $|\psi_{\mathrm{T}}\rangle$ is an arbitrary linear combination of the tranverse polarisation states and $\mathcal{C}$ is an arbitrary constant. Now note that states of different polarizations are *orthogonal* by definition. Hence, $|\phi\rangle$, defined with $(\hat{a}^{\dagger 0}_{\mathbf{p}} - \hat{a}^{\dagger 3}_{\mathbf{p}}) |\Omega\rangle$ is orthogonal to the tranverse polarisation states which are made up of $\hat{a}^{\dagger 1}_{\mathbf{p}}$ and $\hat{a}^{\dagger 2}_{\mathbf{p}}$. Also, we have

$$\langle\phi|\phi\rangle = \langle\Omega| (\hat{a}^0_{\mathbf{p}} - \hat{a}^3_{\mathbf{p}})(\hat{a}^{\dagger 0}_{\mathbf{p}} - \hat{a}^{\dagger 3}_{\mathbf{p}}) |\Omega\rangle = \langle\Omega| (\hat{a}^0_{\mathbf{p}}\hat{a}^{\dagger 0}_{\mathbf{p}} + \hat{a}^3_{\mathbf{p}}\hat{a}^{\dagger 3}_{\mathbf{p}}) |\Omega\rangle$$

Using the commutation relations we can thus write this as:

$$\langle\Omega| (\hat{a}^0_{\mathbf{p}}\hat{a}^{\dagger 0}_{\mathbf{p}} + \hat{a}^3_{\mathbf{p}}\hat{a}^{\dagger 3}_{\mathbf{p}}) |\Omega\rangle = \langle\Omega| \left[\hat{a}^0_{\mathbf{p}}, \hat{a}^{\dagger 0}_{\mathbf{p}}\right] + \left[\hat{a}^3_{\mathbf{p}}, \hat{a}^{\dagger 3}_{\mathbf{p}}\right] |\Omega\rangle = 0$$

where in the last step, due to the difference in the metric signature, the inside term vanishes.

Thus we see that $|\phi\rangle$ is orthogonal to all physical states, since it is orthogonal to the tranverse polarisation states and also to itself. Thus the inner product between $|\psi\rangle$ and any other state will be the same as the inner product between $|\psi_{\mathrm{T}}\rangle$ and that state.

We are kind of getting a feeling that only the transverse polarisation states might matter physically. Let us see what happens when $|\psi\rangle$ acts on the Hamiltonian. From Eq. 39 we can separate out the summation and find the matrix element of the Hamiltonian between two physical states

$$\langle\psi_1| H |\psi_2\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}| \langle\psi| \left(\hat{a}^{\dagger 3}_{\mathbf{p}}\hat{a}^3_{\mathbf{p}} - \hat{a}^{\dagger 0}_{\mathbf{p}}\hat{a}^0_{\mathbf{p}}\right) + (\hat{a}^{\dagger 1}_{\mathbf{p}}\hat{a}^1_{\mathbf{p}} + \hat{a}^{\dagger 2}_{\mathbf{p}}\hat{a}^2_{\mathbf{p}}) |\psi\rangle$$

We can check that applying the Gupta-Bleuer condition on any state $|\gamma\rangle$ we get

$$(\partial_\mu A^\mu)^+ |\gamma\rangle = (\hat{a}^0_{\mathbf{p}} - \hat{a}^3_{\mathbf{p}}) |\gamma\rangle = 0 \implies \boxed{\hat{a}^0_{\mathbf{p}} |\gamma\rangle = \hat{a}^3_{\mathbf{p}} |\gamma\rangle} \quad \text{for any physical state}$$

Using this on the matrix element for the Hamiltonian above, we get that for the non-transverse part,

$$\langle\psi| \hat{a}^{\dagger 3}_{\mathbf{p}} \hat{a}^3_{\mathbf{p}} |\psi\rangle - \langle\psi| \hat{a}^{\dagger 0}_{\mathbf{p}} \hat{a}^0_{\mathbf{p}} |\psi\rangle = \langle\psi| \hat{a}^{\dagger 3}_{\mathbf{p}} \hat{a}^0_{\mathbf{p}} |\psi\rangle - \langle\psi| \hat{a}^{\dagger 0}_{\mathbf{p}} \hat{a}^0_{\mathbf{p}} |\psi\rangle = \langle\psi| \hat{a}^{\dagger 0}_{\mathbf{p}} \hat{a}^0_{\mathbf{p}} |\psi\rangle - \langle\psi| \hat{a}^{\dagger 0}_{\mathbf{p}} \hat{a}^0_{\mathbf{p}} |\psi\rangle = 0$$

Thus, the matrix element of the Hamiltonian for any two physical states become

$$\langle\psi_1| H |\psi_2\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}| \langle\psi| (\hat{a}^{\dagger 1}_{\mathbf{p}} \hat{a}^1_{\mathbf{p}} + \hat{a}^{\dagger 2}_{\mathbf{p}} \hat{a}^2_{\mathbf{p}}) |\psi\rangle$$

The Hamiltonian acts only on the transverse polarizations in the physical subspace and thus, the contribution to energy and momentum comes purely from the tranverse part and is independent of the logitudinal and time-like polarisation part, $\mathcal{C}|\phi\rangle$.

In conclusion, the states $|\psi\rangle = |\psi_{\mathrm{T}}\rangle + \mathcal{C}|\phi\rangle$ and $|\psi_{\mathrm{T}}\rangle$ have the exact same energy and other physical properties and have the same scalar product with all other physical states and are therefore *physically indistinguishable.*

To be a bit more formal, we can specify an equivalence relation $\sim_{\mathrm{R}}$ such that

$$|\gamma\rangle \sim_{\mathrm{R}} |\theta\rangle \iff |\gamma\rangle = |\theta\rangle + \mathcal{C}|\phi\rangle$$

where $|\phi\rangle$ is what was defined above and $\mathcal{C}$ is some constant. Then photons are represented by the equivalence classes of this relation, that is

$$[|\psi_{\mathrm{T}}\rangle]_{\mathrm{R}} := \{|\theta\rangle \mid \ |\theta\rangle \sim_{\mathrm{R}} |\psi_{\mathrm{T}}\rangle\}$$

where we had put the representative of the class as the purely transverse state. Therefore, the photon is described by *two transverse degrees of freedom* as it should be and the generic multiparticle state is obtained tensoring this physical one-particle state.

Lecture 28: Somethings about Propagators

▶ Propagator for Dirac Field

Recall the quantisation for the Dirac field that we had stated. Now, if we move to the Heisenberg picture, we have the field operators as

$$\Psi(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(\hat{a}^s_{\mathbf{p}} u^s(\mathbf{p}) e^{-ip\cdot x} + \hat{b}^{s\dagger}_{\mathbf{p}} v^s(\mathbf{p}) e^{ip\cdot x}\right)$$
$$\overline{\Psi}(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_r \left(\hat{a}^{r\dagger}_{\mathbf{p}} \bar{u}^r(\mathbf{p}) e^{ip\cdot x} + \hat{b}^r_{\mathbf{p}} \bar{v}^r(\mathbf{p}) e^{-ip\cdot x}\right)$$

with the following anti-commutation relations between the annihilation and creation operators,

$$\{\hat{a}^r_{\mathbf{p}}, \hat{a}^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \qquad \{\hat{b}^r_{\mathbf{p}}, \hat{b}^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q})$$

Let us find the anti-commutator between the field and the adjoint field defined by

$$i\mathcal{S}_{\alpha\beta}(x, y) = \{\Psi_\alpha(x), \overline{\Psi}_\beta(y)\}$$

We substitute the expression for the fields in the Heisenberg picture and calculate the anti-commutator

$$\begin{aligned} iS(x,y)_{\alpha,\beta} &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{\sqrt{4E_\mathbf{p}E_\mathbf{q}}}\sum_{r,s}\left[\{\hat{a}^s_\mathbf{p},\hat{a}^{r\dagger}_\mathbf{q}\}u^s_\alpha(\mathbf{p})\bar{u}^r_\beta(\mathbf{q})e^{-i(p\cdot x-q\cdot y)} + \{\hat{b}^{s\dagger}_\mathbf{p},\hat{b}^r_\mathbf{p}\}v^s_\alpha(\mathbf{p})\bar{v}^r_\beta(\mathbf{q})e^{-i(q\cdot y-p\cdot x)}\right] \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\mathrm{d}^3\mathbf{q}\frac{1}{\sqrt{4E_\mathbf{p}E_\mathbf{q}}}\delta^{(3)}(\mathbf{p}-\mathbf{q})\sum_s\left[u^s_\alpha(\mathbf{p})\bar{u}^s_\beta(\mathbf{q})e^{-i(p\cdot x-q\cdot y)} + v^s_\alpha(\mathbf{p})\bar{v}^s_\beta(\mathbf{q})e^{-i(q\cdot y-p\cdot x)}\right] \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}\sum_s\left[u^s_\alpha(\mathbf{p})\bar{u}^s_\beta(\mathbf{p})e^{-ip\cdot(x-y)} + v^s_\alpha(\mathbf{p})\bar{v}^s_\beta(\mathbf{p})e^{ip\cdot(x-y)}\right] \\ &= \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}\left[(\not{p}+m)_{\alpha\beta}e^{-ip\cdot(x-y)} + (\not{p}-m)_{\alpha\beta}e^{ip\cdot(x-y)}\right] \\ &= (i\not{\partial}+m)_{\alpha\beta}D(x-y) - (i\not{\partial}+m)_{\alpha\beta}D(y-x) \\ &= (i\not{\partial}+m)_{\alpha\beta}\triangle(x-y) \end{aligned}$$

where we have used the anti-commutation relations above, also the fact that anti-commutators are symmetric, the completion relations in Eq. ??, Eq. 5 for the expresison of $D(x-y)$ and finally the fact that

$$\partial_\mu e^{\pm ip\cdot(x-y)} = \pm ip_\mu e^{\pm ip\cdot(x-y)} \implies p_\mu e^{\pm ip\cdot(x-y)} = \mp i\partial_\mu e^{\pm ip\cdot(x-y)}$$

In the bosonic case, we saw that $\triangle(x-y)=0$ for space-like intervals, now we get $S(x,y)\equiv S(x-y)\sim\{\Psi_\alpha(x),\overline{\Psi}_\beta(y)\}=0$.

We can define the retarded propagator for the Dirac field using the anti-commutation,

$$\mathcal{S}_\mathrm{R}(x-y) = \Theta(x^0-y^0)\{\Psi(x),\overline{\Psi}(y)\} \equiv (i\not{\partial}+m)D_\mathrm{R}(x-y)\mathbb{1}$$

where $D_\mathrm{R}(x-y)$ is the retarded propagator for the scalar field. It is indeed a Green's function for the Dirac equation since,

$$(i\not{\partial}-m)\mathcal{S}_\mathrm{R}(x-y) = (i\not{\partial}-m)(i\not{\partial}+m)D_\mathrm{R}(x-y) = -(\not{\partial}\not{\partial}+m^2)D_\mathrm{R}(x-y) = i\delta^{(4)}(x-y)\mathbb{1}$$

where we have used $D_\mathrm{R}(x-y)$ is the Green's function for the scalar field and the on-shell condition is satisfied for $\not{\partial}\not{\partial}=\partial^2+m^2$, that is $(\partial^2+m^2)D_\mathrm{R}(x-y)=-i\delta^{(4)}(x-y)$.

If we expand $\mathcal{S}_\mathrm{R}(x-y)$ in the momentum space and use the above equation, we have

$$\begin{aligned} i\delta^{(4)}(x-y) &= (i\not{\partial}-m)\int\frac{\mathrm{d}^4p}{(2\pi)^4}\,\widetilde{\mathcal{S}}_\mathrm{R}(p)e^{-ip\cdot(x-y)} \\ &= \int\frac{\mathrm{d}^4p}{(2\pi)^4}\widetilde{\mathcal{S}}_\mathrm{R}(p)(\not{p}-m)e^{-ip\cdot(x-y)} \end{aligned}$$

Let Dirac delta in the LHS can be expanded as $\int\frac{\mathrm{d}^4p}{(2\pi)^4}\,e^{-ip\cdot(x-y)}$. Then we find that

$$\widetilde{\mathcal{S}}_\mathrm{R}(p)(\not{p}-m) = i\mathbb{1} \quad\xrightarrow{\text{Inverting}}\quad \widetilde{\mathcal{S}}_\mathrm{R}(p) = \frac{i(\not{p}+m)}{p^2-m^2}$$

28.1. Feynman Propagator

We now define the Feynman propagator $\mathcal{S}_\mathrm{F}(x-y)$ for the Dirac field, defined as the time-ordered product

$$\mathcal{S}_\mathrm{F}(x-y) := \langle\Omega|\,\mathcal{T}\Psi(x)\overline{\Psi}(y)\,|\Omega\rangle = \begin{cases}\langle\Omega|\,\Psi(x)\overline{\Psi}(y)\,|\Omega\rangle & x^0>y^0 \\ -\langle\Omega|\,\Psi(y)\overline{\Psi}(x)\,|\Omega\rangle & x^0<y^0\end{cases}$$

The negative sign in the definition is necessary for Lorentz invariance. For two space-like points, no signal or causal influence can travel from one point to the other since these are outside each other's lightcones.

Thus, there should be no invariant way to determine whether $x^0>y^0$ or $x^0<y^0$. As the anti-commutators vanish for space-like separations, the definitions thus become equal for both cases, when

two points are space-like separated. In the bosonic case, the commutators vanished, hence no minus sign was needed.

Similar to the retarded propagator, the Feynman propagator can be expressed as

$$\mathcal{S}_{\mathrm{F}}(x-y) = i\int \frac{\mathrm{d}^4p}{(2\pi)^4}\ e^{-ip\cdot(x-y)}\frac{\not{p}+m}{p^2-m^2+i\epsilon}$$

▶ Propagator for Gauge Field

We will now consider the propagator for the gauge fields that we had been studied for free electromagnetic theory, specifically for photons. Consider the basic Lagrangian for that, without the gauge fixing condition,

$$\mathcal{L}_{\mathrm{EM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

From Eq. 35 we can write the EoM in terms of the gauge field

$$\partial_\mu\partial^\mu A^\nu - \partial_\mu\partial^\nu A^\nu = 0 \implies \partial_\mu\partial^\mu A^\nu - \partial^\mu\partial^\nu A_\mu = 0 \implies \boxed{(\eta^{\sigma\nu}\partial^\mu\partial_\mu - \partial^\sigma\partial^\nu)A_\sigma = 0}$$

In order to find the propagator, we can now try to find a Green's function for this equation, say $G_{\rho\nu}(x-y)$ satisfying

$$(\eta^{\sigma\nu}\partial^\mu\partial_\mu - \partial^\sigma\partial^\nu)G_{\rho\nu}(x-y) = i\delta^{(4)}(x-y)\delta^\sigma{}_\rho$$

Now expand the Green's function using its Fourier modes and then substitute in the equation above

$$G_{\rho\nu}(x-y) = \int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ip\cdot(x-y)}\widetilde{G}_{\rho\nu}(p) \longrightarrow \int\frac{\mathrm{d}^4p}{(2\pi)^4}\widetilde{G}_{\rho\nu}(p)(\eta^{\sigma\nu}\partial^\mu\partial_\mu - \partial^\sigma\partial^\nu)e^{-ip\cdot(x-y)} = i\delta^\sigma{}_\rho\int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ip\cdot(x-y)}$$

$$\implies \int\frac{\mathrm{d}^4p}{(2\pi)^4}\widetilde{G}_{\rho\nu}(p)(-\eta^{\sigma\nu}p^2 + p^\sigma p^\nu)e^{-ip\cdot(x-y)} = \int\frac{\mathrm{d}^4p}{(2\pi)^4}i\delta^\sigma{}_\rho e^{-ip\cdot(x-y)}$$

Comparing both sides, we get the relation

$$(-\eta^{\sigma\nu}p^2 + p^\sigma p^\nu)\widetilde{G}_{\rho\nu}(p) = i\delta^\sigma{}_\rho$$

which implies that $\widetilde{G}_{\rho\nu}(p)$ could be an inverse of the operator $(-\eta^{\sigma\nu}p^2 + p^\sigma p^\nu)$. However, note that

$$(-\eta^{\sigma\nu}p^2 + p^\sigma p^\nu)p_\nu = (-p^2p^\sigma + p^2p^\sigma) = 0$$

and hence 0 is one of the eigenvalues of this operator. This implies that this operator is singular since its determinant becomes zero as the determinant is written as the product of the eignevalues. As a result, the operator cannot be inverted and the Green's function cannot be stated to be the inverse of the operator and its existence becomes questionable!

To solve this problem, we fall back to our Gauge fixing Lagrangian

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

which has the corresponding EoM from Eq. 36. Finding the same Green's function in momentum space, we can find that a factor of $(1 - {}^1\!/\!_\xi)$ comes into the second term of the operator which saves it from being singular and hence, it now becomes invertible.

To construct the inverse matrix, we make a general symmetric ansatz

$$\widetilde{G}_{\rho\nu}(p) = \mathcal{A}(p)\eta_{\rho\nu} + \mathcal{B}(p)p_\rho p_\nu$$

We note substitute this in the modified inverse equation

$$\begin{aligned} i\delta^\sigma{}_\rho &= \left(-\eta^{\sigma\nu}p^2 + \left(1-\frac{1}{\xi}\right)p^\sigma p^\nu\right)(\mathcal{A}(p)\eta_{\rho\nu} + \mathcal{B}(p)p_\rho p_\nu) \\ &= -\mathcal{A}(p)p^2\delta^\sigma{}_\rho - \mathcal{B}(p)p_\rho p^\sigma p^2 + \left(1-\frac{1}{\xi}\right)p_\rho p^\sigma\mathcal{A}(p) + \mathcal{B}(p)\left(1-\frac{1}{\xi}\right)p^2p^\rho p_\sigma \end{aligned}$$

Comparing both sides, we get $\mathcal{A}(p) = -i/p^2$ and

$$-\mathcal{B}(p)\frac{p^2}{\xi} + \left(1 - \frac{1}{\xi}\right)\mathcal{A}(p) = 0 \implies \mathcal{B}(p) = -\frac{i}{p^4}(\xi - 1)$$

Thus the Green's function in the momentum space is found as

$$\widetilde{G}_{\rho\nu}(p) = -\frac{i}{p^2}\left[\eta_{\rho\nu} + \frac{\xi - 1}{p^2}p_\rho p_\nu\right]$$

Then choosing different pole prescriptions, we can find the different propagators viz. retarded, advanced, Feynman... For example, the Feynman propagator for the photon field is given as

$$D_{\mathrm{F}}(x - y) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4}\left\{-\frac{i}{p^2 + i\varepsilon}\left[\eta_{\rho\nu} + \frac{\xi - 1}{(p^2 + i\varepsilon)^2}p_\rho p_\nu\right]\right\}e^{-ip\cdot(x-y)}$$

Lecture 29: Introducing Interactions

So far we had been discussing the free field theory, where there was no interactions between the particles. We had a hint of interactions emerging when we tried to impose local gauge symmetry on the Lagrangian, when two-point and four-point interaction terms automatically emerged by the use of covariant derivatives.

Emergence of interesting phenomenon only happens in presence of *interactions* 😉 and thus we need to add some extra terms to the Lagrangian, beyond the bilinear terms. The problem becomes intractable analytically, however, if the strength of the interaction is small, we often fall back to *perturbation theory* to find some approximate solutions.

29.1. Allowed Interactions

Consider the Klein-Gordon Lagrangian as an example

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

Intuitively we expect that an infinite number of interactions can be added to the system, however not all type of interactions are allowed or feasible for the theory. To understand the allowed interactions, we need to see some subtleties.

Note that the action has dimensions of $\hbar$ but we are working in natural units where $\hbar$ is dimensionless and hence, dimension of action $[\mathcal{S}] = 0$. Moreover, from sec 2 we know that

$$[L] = [M^{-1}] \implies [\mathrm{d}^4 x] = [M^{-4}]$$

Then from the fact that $\mathcal{S} = \int \mathrm{d}^4 x\ \mathcal{L}$, we obtain $[\mathcal{L}] = [M^4]$. Note that the partial derivative is kinda inverse length and $[\partial_\mu] = [M]$. Then the typical term of the Lagrangian, $[m^2\phi^2] = [M^4] \implies [\phi] = [M]$. Thus if we want to add some interaction of the form $\lambda_n \phi^n$ to the Lagrangian, $[\lambda_n] = [M^{4-n}]$

If $n \neq 4$ then λ_n has non-negative mass dimension, then the theory is said to be *renormalisable* (which relates to manipulating some divergences appearing in integrals) and *non-renormalisable* otherwise (which become irrelevant for us)!

We will start with our discussion for $n = 4$, that is the Lagrangian is of the form

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

29.2. The Interaction Picture

Working with Hamiltonians is somewhat easier now, hence we define the Hamiltonian in terms of the free Hamiltonian and the interacting Hamiltonian

$$\mathrm{H} = \mathrm{H}_0 + \mathrm{H}_{\text{int}}$$

The field configuration for the Hamiltonian, $\phi(x)$ will definitely not be a simple plane wave. In order to expand perturbatively we need to relate the field ϕ to some ϕ_{I} whose evolution is determined by H_0 alone. We define such a field, called the *interaction picture field* by the following evolution

$$\phi_{\text{I}}(t, \mathbf{x}) = e^{i\mathrm{H}_0(t-t_0)}\phi_{\text{I}}(t_0, \mathbf{x})e^{-i\mathrm{H}_0(t-t_0)}$$

where t_0 is our *reference time*. The interaction field evolves with the free Hamiltonian and thus can be safely expanded as

$$\phi_{\text{I}}(t, \mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(\hat{a}_{\mathbf{p}}e^{-ip\cdot x} + \hat{a}_{\mathbf{p}}^{\dagger}e^{ip\cdot x}\right)$$

with the usual creation and annihilation operators. We want to express the full Heisenberg field in terms of ϕ_{I}. Suppose at $t = t_0$ both the interaction field and the Heisenberg field are the same. We now use the evolution for the Heiserberg fields (which evolve with the total Hamiltonian H)

$$\begin{aligned}
\phi(t, \mathbf{x}) &= e^{i\mathrm{H}\tau}\phi(t_0, \mathbf{x})e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}\tau}e^{-i\mathrm{H}_0\tau}e^{i\mathrm{H}_0\tau}\phi(t_0, \mathbf{x})e^{-i\mathrm{H}_0\tau}e^{i\mathrm{H}_0\tau}e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}\tau}e^{-i\mathrm{H}_0\tau}e^{i\mathrm{H}_0\tau}\phi(t_0, \mathbf{x})e^{-i\mathrm{H}_0\tau}e^{i\mathrm{H}_0\tau}e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}\tau}e^{-i\mathrm{H}_0\tau}\phi_{\text{I}}(t, \mathbf{x})e^{i\mathrm{H}_0\tau}e^{-i\mathrm{H}\tau}
\end{aligned}$$

where we have define $\tau := (t - t_0)$. In the second step, we have introduced the identity operator and then used the evolution of the interaction field from above.

It is useful to define the unitary operator $U(t, t_0) := e^{i\mathrm{H}_0\tau}e^{-i\mathrm{H}\tau}$ such that we can write

$$\boxed{\phi(t, \mathbf{x}) = U^{\dagger}(t, t_0)\phi_{\text{I}}(t, \mathbf{x})U(t, t_0)}$$

We need to find some evolution equation for this unitary operator which we find by its time-derivative

$$\begin{aligned}
i\frac{\partial U}{\partial t} = -\mathrm{H}_0e^{i\mathrm{H}_0\tau}e^{-i\mathrm{H}\tau} + e^{i\mathrm{H}_0\tau}\mathrm{H}e^{-i\mathrm{H}\tau} &= -e^{i\mathrm{H}_0\tau}\mathrm{H}_0e^{-i\mathrm{H}\tau} + e^{i\mathrm{H}_0\tau}\mathrm{H}e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}_0\tau}(\mathrm{H} - \mathrm{H}_0)e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}_0\tau}H_{\text{int}}e^{-i\mathrm{H}\tau} \\
&= e^{i\mathrm{H}_0\tau}\mathrm{H}_{\text{int}}e^{-i\mathrm{H}_0}e^{i\mathrm{H}_0}e^{-i\mathrm{H}\tau}
\end{aligned}$$

Defining the *interaction picture Hamiltonian* as $\mathrm{H}_{\text{I}}(t) = e^{i\mathrm{H}_0\tau}\mathrm{H}_{\text{int}}e^{-i\mathrm{H}_0\tau}$ we get the evolution equation as

$$\boxed{i\frac{\partial U}{\partial t} = \mathrm{H}_{\text{I}}(t)U(t, t_0)} \tag{40}$$

29.2.1. Dyson Series

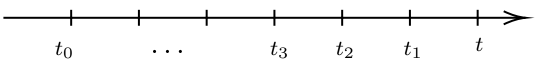
We now try to find a solution for the above differential equation of the unitary operator U. First note the boundary condition, that $U(t_0, t_0) = \mathbb{1}$. Then the solution of Eq. 40 can be iteratively written as

$$\begin{aligned}
U(t,t_0) &= \mathbb{1} - i \int_{t_0}^{t} \mathrm{d}t_1 \; \mathrm{H_I}(t_1) U(t_1, t_0) \\
&= \mathbb{1} - i \int_{t_0}^{t} \mathrm{d}t_1 \; \mathrm{H_I}(t_1) \left[\mathbb{1} - i \int_{t_0}^{t_1} \mathrm{d}t_2 \; \mathrm{H_I}(t_2) U(t_2, t_0) \right] \\
&= \mathbb{1} - i \int_{t_0}^{t} \mathrm{d}t_1 \; \mathrm{H_I}(t_1) + (-i)^2 \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \; \mathrm{H_I}(t_1) \mathrm{H_I}(t_2) U(t_2, t_0) \\
&= \mathbb{1} - i \int_{t_0}^{t} \mathrm{d}t_1 \; \mathrm{H_I}(t_1) + (-i)^2 \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \; \mathrm{H_I}(t_1) \mathrm{H_I}(t_2) \left[\mathbb{1} - i \int_{t_0}^{t_2} \mathrm{d}t_3 \; \mathrm{H_I}(t_3) U(t_3, t_0) \right] \\
&= \mathbb{1} - i \int_{t_0}^{t} \mathrm{d}t_1 \; \mathrm{H_I}(t_1) + (-i)^2 \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \; \mathrm{H_I}(t_1) \mathrm{H_I}(t_2) + \\
&\qquad\qquad (-i)^3 \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \int_{t_0}^{t_2} \mathrm{d}t_3 \; \mathrm{H_I}(t_1) \mathrm{H_I}(t_2) \mathrm{H_I}(t_3) U(t_3, t_0)
\end{aligned}$$

Then the general solution can be written as

$$U(t,t_0) = \mathbb{1} + \sum_{n=1}^{\infty} (-i)^n \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \ldots \int_{t_0}^{t_{\mathrm{n-1}}} \mathrm{d}t_{\mathrm{n}} \left(\prod_{j=1}^{n} \mathrm{H_I}(t_{\mathrm{j}}) \right)$$

Note that in the above expression, from the integration limits we can observe that $t_{\mathrm{i}} \in (t_0, t_{\mathrm{i+1}})$ and hence $t_{\mathrm{i+1}} < t_{\mathrm{i}}$



Hence a natural time-ordering appears since the Hamiltonians inside the integrals are arranged in a time-ordered fashion. Then using the time-ordering operator $\mathcal{T}$ (which acts on a 'string' of operators at different times and arranges them with greater times appearing on the left) we can write

$$U(t,t_0) = \mathbb{1} + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t} \mathrm{d}t_2 \ldots \int_{t_0}^{t} \mathrm{d}t_{\mathrm{n}} \; \mathcal{T} \left\{ \prod_{j=1}^{n} \mathrm{H_I}(t_{\mathrm{j}}) \right\}$$

We observe that in the above integrals we can permute between t_i since there is no ordering. The region of integral in the above expression can be broken into $n!$ sub-regions, depending on the permutations of the times, hence we divide by $n!$ for this purpose. We symbolically define the integral as the *time-ordered exponential*

$$U(t,t_0) = \mathcal{T} \left\{ \exp \left[-i \int_{t_0}^{t} \mathrm{d}t' \mathrm{H_I}(t') \right] \right\} \tag{41}$$

This compact form of the operator is called the Dyson series and is extremely hard to compute fully. Computations are often done keeping only few terms and the equation is kind of a formality.

Fun Fact:
The following are two important properties of the unitary operator U:

- $U^\dagger(t_2, t_0) = U(t_0, t_2)$
- $U(t_1, t_0)U(t_0, t_2) = U(t_1, t_2)$

Proof. Let us try to prove these properties one by one.

▷ For the first property, note that taking dagger of Eq. 41 changes $i \to -i$ and this extra minus sign is then used to interchange the limits of the integral. Hence in the exponential we have $-i\int_{t_2}^{t_0}$, thus proved.

▷ For the second property, let us define $\widetilde{U}(t) = U(t, t_0)U(t_0, t_2)$. Then we have

$$i\frac{\partial \widetilde{U}}{\partial t} = i\left(\frac{\partial}{\partial t}U(t, t_0)\right)U(t_0, t_2) = \mathrm{H_I}(t, t_0)U(t, t_0)U(t_0, t_2) = \mathrm{H_I}(t)\widetilde{U}(t)$$

Thus, $\widetilde{U}$ satisfies the same differential equation as $U(t, t_0)$ and we have $\widetilde{U}(t_2) = U(t_2, t_0)U(t_0, t_2) = \mathbb{1}$ using the first property and unitarity of U.

Thus, $\widetilde{U}(t)$ satisfies the same equation and same initial conditon as $U(t, t_2)$ (the reference time is replaced, $t_0 \to t_2$). Thus by the uniqueness of the solution of first-order linear differential equation $\widetilde{U}(t) = U(t, t_2)$ and hence proved.

29.3. Ground States

The ground state of the Hamiltonian is very special since it acts as a stable-reference state of the system. It kind of denotes the zero-temperature QFT(which is simpler since at higher temperature, the system stays in a *thermal ensemble*). Indeed we had calculated the propagators as the *vacuum* expectation of some shitty quantities.

Thus characterising ground states becomes very important.We had already said a lot of bs about the ground states of the free theory $|\Omega\rangle$, however, in presence of interactions, the ground state of the Hamiltonian also changes to some $|\widetilde{\Omega}\rangle$ and is unknown to us.

Let the full Hamiltonian be characterised by the eigenpairs $\{|n\rangle, \mathcal{E}_n\}$ where for simplicity we assume these to be countable(when actually these form a continuum) and non-degenerate.

$$\begin{aligned} e^{-i\mathrm{H}\tau}|\Omega\rangle &= e^{-i\mathrm{H}\tau}\left(\sum_{n\neq 0}|n\rangle\langle n| + |\widetilde{\Omega}\rangle\langle\widetilde{\Omega}|\right)|\Omega\rangle \\ &= \sum_{n\neq 0} e^{-i\mathcal{E}_\mathrm{n}\tau}\langle n|\Omega\rangle|n\rangle \quad + e^{-i\mathcal{E}_0\tau}\langle\widetilde{\Omega}|\Omega\rangle|\widetilde{\Omega}\rangle \end{aligned}$$

Note the $\mathcal{E}_\mathrm{n} > \mathcal{E}_0$ for all n since it is the ground (lowest energy) state of the interacting theory. Then we can use a trick, sending τ to infinity in a *slightly imaginary* manner, such that the only contribution is for $n = 0$ (the other terms become exponentially decaying). Thus we obtain

$$\lim_{\tau\to\infty(1-i\epsilon)} e^{-i\mathrm{H}\tau}|\Omega\rangle = \lim_{\tau\to\infty(1-i\epsilon)}\left[e^{-i\mathcal{E}_0\tau}\langle\widetilde{\Omega}|\Omega\rangle\right]|\widetilde{\Omega}\rangle$$

Thus we obtain an expression for the vacuum of the interacting theory as

$$|\widetilde{\Omega}\rangle = \frac{\lim_{\tau\to\infty(1-i\epsilon)} e^{-i\mathrm{H}\tau}|\Omega\rangle}{\lim_{\tau\to\infty(1-i\epsilon)}\left[e^{-i\mathcal{E}_0\tau}\langle\widetilde{\Omega}|\Omega\rangle\right]}$$

In the expression above, we assume that there is some finite overlap between the vacuums of the free and interacting theory. If the overlap is zero, we can say that the interaction is in no way ‘small’ and perturbative analysis becomes invalid!

Now, note that τ goes to infinite and it would not matter if we shift this by some finite amount, $\tau \to \tau + t_0$

$$
\begin{aligned}
|\widetilde{\Omega}\rangle &= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathrm{H}(\tau+t_0)}|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(\tau+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathrm{H}(-(-\tau)+t_0)}|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(-(-\tau)+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]}
\end{aligned}
$$

Since $|\Omega\rangle$ is the ground-state of the free Hamiltonian, $\mathrm{H}_0|\Omega\rangle = 0 \implies e^{i\mathrm{H}_0(\tau+t_0)}|\Omega\rangle = |\Omega\rangle$. We can introduce this in the expression above

$$
\begin{aligned}
|\widetilde{\Omega}\rangle &= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathrm{H}(-(-\tau)+t_0)}e^{i\mathrm{H}_0(\tau+t_0)}|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(-(-\tau)+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{i\mathrm{H}((-\tau)-t_0)}e^{-i\mathrm{H}_0(-\tau-t_0)}|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(-(-\tau)+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} U^\dagger(-\tau,t_0)|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(-(-\tau)+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} U(t_0,-\tau)|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(-(-\tau)+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]}
\end{aligned}
$$

Intuitively, this says that upto some numerical factor, the vacuum of the interacting theory is obtained by evolving the vacuum of the free theory from $-\tau$ to t_0 where τ is some very very large time. To obtain $\langle\widetilde{\Omega}|$ we start with $e^{i\mathrm{H}t}|\Omega\rangle$ instead. Expanding as before we see

$$
e^{i\mathrm{H}\tau}|\Omega\rangle = \sum_{n\neq 0} e^{i\mathcal{E}_n\tau}\langle n|\Omega\rangle|n\rangle \quad + e^{i\mathcal{E}_0\tau}\langle\widetilde{\Omega}|\Omega\rangle|\widetilde{\Omega}\rangle
$$

Taking Hermitian conjugate both sides, we obtain

$$
\langle\Omega|e^{-i\mathrm{H}\tau} = \sum_{n\neq 0}\langle n|e^{-i\mathcal{E}_n\tau}\langle\Omega|n\rangle \quad + \langle\widetilde{\Omega}|e^{-i\mathcal{E}_0\tau}\langle\Omega|\widetilde{\Omega}\rangle
$$

By the same previous logic, taking τ to infinity in a slightly imaginary manner we get

$$
\begin{aligned}
\langle\widetilde{\Omega}| &= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|e^{-i\mathrm{H}\tau}}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0\tau}\langle\Omega|\widetilde{\Omega}\rangle} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|e^{-i\mathrm{H}(\tau-t_0)}}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0(\tau-t_0)}\langle\Omega|\widetilde{\Omega}\rangle} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|e^{i\mathrm{H}_0(\tau-t_0)}e^{-i\mathrm{H}(\tau-t_0)}}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0(\tau-t_0)}\langle\Omega|\widetilde{\Omega}\rangle} \\
&= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|U(\tau,t_0)}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0(\tau-t_0)}\langle\Omega|\widetilde{\Omega}\rangle}
\end{aligned}
$$

Thus we had found the expression of the vacuum for the interacting theory

$$\boxed{|\widetilde{\Omega}\rangle = \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} U(t_0,-\tau)|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} \left[e^{-i\mathcal{E}_0(\tau+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]}} \qquad \boxed{\langle\widetilde{\Omega}| = \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|\,U(\tau,t_0)}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0(\tau-t_0)}\langle\Omega|\widetilde{\Omega}\rangle}} \tag{42}$$

29.4. Correlation Functions

For interacting theories, we will often have to calculate the n-point correlation function of the form $\langle\widetilde{\Omega}|\,\phi(x_1)\phi(x_2)\ldots\phi(x_n)\,|\widetilde{\Omega}\rangle$. Consider a time-ordered two-point correlation function between $\phi(x)$ and $\phi(y)$ such that $x^0 > y^0$.

$$\begin{aligned}
\langle\widetilde{\Omega}|\,\phi(x)\phi(y)\,|\widetilde{\Omega}\rangle &= \frac{\lim\limits_{\tau\to\infty(1-i\epsilon)} \langle\Omega|\,U(\tau,t_0)\phi(x)\phi(y)U(t_0,-\tau)|\Omega\rangle}{\lim\limits_{\tau\to\infty(1-i\epsilon)} e^{-i\mathcal{E}_0(\tau-t_0)}\langle\Omega|\widetilde{\Omega}\rangle\left[e^{-i\mathcal{E}_0(\tau+t_0)}\langle\widetilde{\Omega}|\Omega\rangle\right]} \\
&= \lim_{\tau\to\infty(1-i\epsilon)}\left[e^{-2i\mathcal{E}_0\tau}|\,\langle\Omega|\widetilde{\Omega}\rangle\,|^2\right]^{-1}\langle\Omega|\,U(\tau,t_0)\left\{U^\dagger(x^0,t_0)\phi_{\mathrm{I}}(x)U(x^0,t_0)\right\} \\
&\qquad\times\left\{U^\dagger(y^0,t_0)\phi_{\mathrm{I}}(y)U(y^0,t_0)\right\}U(t_0,-\tau)|\Omega\rangle \\
&= \lim_{\tau\to\infty(1-i\epsilon)}\left[e^{-2i\mathcal{E}_0\tau}|\,\langle\Omega|\widetilde{\Omega}\rangle\,|^2\right]^{-1}\langle\Omega|\,U(\tau,x^0)\phi_{\mathrm{I}}(x)U(x^0,y^0)\phi_{\mathrm{I}}(y)U(y^0,-\tau)|\Omega\rangle
\end{aligned}$$

Now consider that the vacuum states are normalised, $\langle\widetilde{\Omega}|\widetilde{\Omega}\rangle = 1$ and then the two-point correlation becomes

$$\begin{aligned}
\frac{\langle\widetilde{\Omega}|\,\phi(x)\phi(y)\,|\widetilde{\Omega}\rangle}{\langle\widetilde{\Omega}|\widetilde{\Omega}\rangle} &= \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,U(\tau,x^0)\phi_{\mathrm{I}}(x)U(x^0,y^0)\phi_{\mathrm{I}}(y)U(y^0,-\tau)|\Omega\rangle}{\langle\Omega|\,U(\tau,t_0)U(t_0,-\tau)\,|\Omega\rangle} \\
&= \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,U(\tau,x^0)\phi_{\mathrm{I}}(x)U(x^0,y^0)\phi_{\mathrm{I}}(y)U(y^0,-\tau)|\Omega\rangle}{\langle\Omega|\,U(\tau,-\tau)\,|\Omega\rangle}
\end{aligned}$$

Note that the entire expression above is time-ordered. Then we can write the above expression as

$$\begin{aligned}
\langle\widetilde{\Omega}|\,\mathcal{T}\{\phi(x)\phi(y)\}\,|\widetilde{\Omega}\rangle &= \lim_{\tau\to\infty(1-i\epsilon)} \frac{\mathcal{T}\{\langle\Omega|\,U(\tau,x^0)\phi_{\mathrm{I}}(x)U(x^0,y^0)\phi_{\mathrm{I}}(y)U(y^0,-\tau)|\Omega\rangle\}}{\langle\Omega|\,U(\tau,-\tau)\,|\Omega\rangle} \\
&= \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,\mathcal{T}\{\phi_{\mathrm{I}}(x)\phi_{\mathrm{I}}(y)U(\tau,x^0)U(x^0,y^0)U(y^0,-\tau)\}|\Omega\rangle}{\langle\Omega|\,U(\tau,-\tau)\,|\Omega\rangle} \\
&= \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,\mathcal{T}\{\phi_{\mathrm{I}}(x)\phi_{\mathrm{I}}(y)U(\tau,-\tau)\}|\Omega\rangle}{\langle\Omega|\,U(\tau,-\tau)\,|\Omega\rangle}
\end{aligned}$$

Now using the Dyson series from Eq. 41 we can write the final expression for the two-point correlation function as

$$\langle\widetilde{\Omega}|\,\mathcal{T}\{\phi(x)\phi(y)\}\,|\widetilde{\Omega}\rangle = \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,\mathcal{T}\left\{\phi_{\mathrm{I}}(x)\phi_{\mathrm{I}}(y)\exp\left[-i\int_{-\tau}^{\tau}\mathrm{d}t'\ \mathrm{H}_{\mathrm{I}}(t')\right]\right\}|\Omega\rangle}{\langle\Omega|\,\mathcal{T}\left\{\exp\left[-i\int_{-\tau}^{\tau}\mathrm{d}t'\ \mathrm{H}_{\mathrm{I}}(t')\right]\right\}|\Omega\rangle}$$

Generalising to n-point correlation function we have the expression

$$\boxed{\langle\widetilde{\Omega}|\,\mathcal{T}\{\phi(x_1)\phi(x_2)\ldots\phi(x_n)\}\,|\widetilde{\Omega}\rangle = \lim_{\tau\to\infty(1-i\epsilon)} \frac{\langle\Omega|\,\mathcal{T}\left\{\phi_{\mathrm{I}}(x_1)\phi_{\mathrm{I}}(x_2)\ldots\phi_{\mathrm{I}}(x_n)\exp\left[-i\int_{-\tau}^{\tau}\mathrm{d}t'\ \mathrm{H}_{\mathrm{I}}(t')\right]\right\}|\Omega\rangle}{\langle\Omega|\,\mathcal{T}\left\{\exp\left[-i\int_{-\tau}^{\tau}\mathrm{d}t'\ \mathrm{H}_{\mathrm{I}}(t')\right]\right\}|\Omega\rangle}} \tag{43}$$

The expression tells us how to calculate the correlation function for the interacting theory in terms of the interaction picture fields (which are *free fields* since these evolve with the free Hamiltonian). Also the H_{I} has the same functional dependence on the ϕ_{I} as $\mathrm{H}_{\mathrm{int}}$ has on ϕ and thus it becomes easier. However, this is indeed a brute force method for computation and becomes very cumbersome!

Lecture 30: Wick's Theorem

In the previous section we looked into an introduction to interacting fields and found general expressions for the correlation functions

$$\langle\Omega|\,\mathcal{T}\{\phi(x_1)\phi(x_2)\ldots\phi(x_n)\}\,|\Omega\rangle$$

where we need to find the time-ordered product of free fields. These are efficiently given by a method of computation. As a gross approximation, Wick's theorem tells us how to go from time ordered products to normal ordered products.

Recall from sec 11.2 that normal ordered products meant that all the creation operators are sent to the left. This is particularly useful, since if creation operators are on the left (and annihilation are on the right), then these will kill the vacuum from both sides when expectation is taken, which might simplify our calculations somewhat.

We will use the notation $\phi(x_i) \equiv \phi_i$ and for the Feynman propagator, $D_{\mathrm{F}}(x_i - x_j) \equiv \triangle_{ij}$

30.1. Few calculations

For $n = 1$, $\mathcal{T}\{\phi_1\} = \phi_1$ since there is only one field. For $n = 2$,

$$\mathcal{T}\{\phi_1, \phi_2\} = \Theta(x_1^0 - x_2^0)\phi_1 + \Theta(x_2^0 - x_1^0)\phi_2$$

which we already know, however, let us see this a bit differently. For that, decompose the Heisenberg field $\phi(x)$ into $\phi^+(x) + \phi^-(x)$ where

$$\phi^+(x) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{p}}}\hat{a}_\mathbf{p}e^{-ip\cdot x}$$
$$\phi^-(x) := \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{p}}}\hat{a}_\mathbf{p}^\dagger e^{ip\cdot x}$$

Note the unfortunate notation that + comes with annihilation and − comes with creation operators. Normal ordering is there if minuses (if any) are to the left. Now suppose that we assume the time-order $x_1^0 > x_2^0$ which gives

$$\begin{aligned}\mathcal{T}\{\phi_1, \phi_2\} = \phi_1\phi_2 &= (\phi_1^+ + \phi_1^-)(\phi_2^+ + \phi_2^-)\\ &= \phi_1^+\phi_2^+ + \phi_1^+\phi_2^- + \phi_1^-\phi_2^+ + \phi_1^-\phi_2^-\\ &= \phi_1^+\phi_2^+ + \phi_2^-\phi_1^+ + \phi_1^-\phi_2^+ + \phi_1^-\phi_2^- + \left[\phi_1^+, \phi_2^-\right]\end{aligned}$$

The blue term was not normal ordered and hence in the last line, we made that normal ordered by introducing the commutator. The commutator gives us $\left[\phi_1^+, \phi_2^-\right] = D(x-y) \equiv D_{12}$ where $D(x-y)$ is defined as in sec 15.

As a shorthand, we denote the normal-ordering by $\mathcal{N}\{\phi_1, \phi_2\} \equiv\, :\phi_1\phi_2:$, hence for $x_1^0 > x_2^0$

$$\mathcal{T}\{\phi_1, \phi_2\} = \,:\phi_1\phi_2: + D_{12}$$

If we calculate for the other case, $x_1^0 < x_2^0$ then 1 and 2 are simply exchanged,

$$\mathcal{T}\{\phi_1, \phi_2\} = \,:\phi_1\phi_2: + D_{21}$$

Thus in general we can write the time-ordered product in the following form

$$\begin{aligned}\mathcal{T}\{\phi_1, \phi_2\} = &\ \Theta(x_2^0 - x_1^0)[:\phi_1\phi_2: + D_{21}] + \Theta(x_1^0 - x_2^0)[:\phi_1\phi_2: + D_{12}]\\ &= \,:\phi_1\phi_2: + \Theta(x_2^0 - x_1^0)D_{21} + \Theta(x_1^0 - x_2^0)D_{21}\\ &= \,:\phi_1\phi_2: + \triangle_{12}\end{aligned}$$

We should have been more careful with the above expression. The LHS is an operator, however the Feynman propagator $\triangle_{\text{ij}}$ is just a complex number, hence we should have technically written $\triangle_{12}\mathbb{1}$. We define a *contraction* as

$$\contraction{}{\phi_1}{}{\phi_2}\phi_1\phi_2 := \triangle_{12}\mathbb{1}$$

Then the time-ordered product of two-fields can be written as

$$\mathcal{T}\{\phi_1, \phi_2\} = :\phi_1\phi_2: + \contraction{}{\phi_1}{}{\phi_2}\phi_1\phi_2$$

30.2. The main theorem

Theorem 3 (Wick's Theorem):
The time-ordered pairs of fields is given by

$$\mathcal{T}\{\phi_1, \phi_2 \dots \phi_n\} = \sum_{m=0}^{\lfloor n/2 \rfloor} \sum_{\substack{\text{all } m \\ \text{pairings}}} : \underbrace{\phi_1\phi_2 \dots \phi_n}_{\text{m contractions}} :$$

What the theorem says is that the time-ordered product of n fields is the equal to the sum of operators formed by all possible contractions between the fields.

We now try to prove the theorem but before that let us see an example for $n = 4$. Note that any contraction will remove 2 fields and hence the maximum number of contractions possible is $\lfloor n/2 \rfloor$. If n is odd, then one field is always uncontracted.

Applying this to four fields, we get

$$\begin{aligned}
\mathcal{T}\{\phi_1, \phi_2, \phi_3, \phi_4\} = {} & :\phi_1\phi_2\phi_3\phi_4: + \contraction{}{\phi_1}{}{\phi_2}\phi_1\phi_2\phi_3\phi_4 + \contraction{}{\phi_1}{\phi_2}{\phi_3}\phi_1\phi_2\phi_3\phi_4 + \contraction{}{\phi_1}{\phi_2\phi_3}{\phi_4}\phi_1\phi_2\phi_3\phi_4 + \contraction{\phi_1}{\phi_2}{}{\phi_3}\phi_1\phi_2\phi_3\phi_4 \\
& \contraction{\phi_1}{\phi_2}{\phi_3}{\phi_4}\phi_1\phi_2\phi_3\phi_4 + \contraction{\phi_1\phi_2}{\phi_3}{}{\phi_4}\phi_1\phi_2\phi_3\phi_4 + \contraction{}{\phi_1}{}{\phi_2}\contraction{\phi_1\phi_2}{\phi_3}{}{\phi_4}\phi_1\phi_2\phi_3\phi_4 + \contraction{}{\phi_1}{\phi_2}{\phi_3}\contraction[2ex]{\phi_1}{\phi_2}{\phi_3}{\phi_4}\phi_1\phi_2\phi_3\phi_4 + \contraction{}{\phi_1}{\phi_2\phi_3}{\phi_4}\contraction[2ex]{\phi_1}{\phi_2}{}{\phi_3}\phi_1\phi_2\phi_3\phi_4
\end{aligned}$$

where the contraction between two fields not adjacent to each other, like say $\contraction{}{\phi_1}{\phi_2}{\phi_3}\phi_1\phi_2\phi_3\phi_4$, means that the Feynman propagator is between those two fields and the other fields are normal-ordered, that is,

$$(\triangle_{13})(:\phi_2\phi_4:)$$

Proof. We have already proved it for $n = 1$ and $n = 2$. Let fields $\phi_1, \dots, \phi_n$ be given and WLOG we can relabel the fields such that these become time-ordered, that is, $x_1^0 \geq x_2^0 \dots \geq x_n^0$ and we claim that Wick's theorem is true for $(n-1)$ fields, $\phi_2, \phi_3 \dots \phi_n$, that is

$$\mathcal{T}\{\phi_2 \dots \phi_n\} = \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \sum_{\substack{\text{all } m \\ \text{pairings}}} : \underbrace{\phi_1\phi_2 \dots \phi_n}_{\text{m contractions}} : \equiv \mathcal{W}(\phi_2 \dots \phi_n)$$

We have to now show that it is true for n fields. Multiply both sides of the above equation by ϕ_1. Since x_1^0 is greater than all other terms, the LHS becomes $\mathcal{T}\{\phi_1\phi_2 \dots \phi_n\}$. Now the RHS becomes

$$\phi_1\mathcal{W}(\phi_2 \dots \phi_n) = (\phi_1^+ + \phi_1^-)\mathcal{W}(\phi_2 \dots \phi_n)$$

The second term is not problematic, since all the terms in $\mathcal{W}$ are already normal ordered, so ϕ_1^- can safely move in to the left, without disturbing the normal-ordering. For ϕ_1^+ we need to send it to the extreme right and thus, we have to invoke the commutator.

$$\phi_1\mathcal{W}(\phi_2 \dots \phi_n) = \phi_1^-\mathcal{W}(\phi_2 \dots \phi_n) + \mathcal{W}(\phi_2 \dots \phi_n)\phi_1^+ + \left[\phi_1^+, \mathcal{W}(\phi_2 \dots \phi_n)\right]$$

The first two terms together contain all possible contractions without ϕ_1 and multiplied with ϕ_1, so it has all possible normal-ordered cases where ϕ_1 is not contracted. Now consider the commutator part (the third term). Note that $\mathcal{W}(\phi_2 \dots \phi_n)$ contains all contractions which do not contain ϕ_1. So it has a general form of some contraction times a normal-ordered product of fields,

$$\text{(contraction)}\left[\phi_1^+, \phi_{i_1}\phi_{i_2}\dots\phi_{i_k}\right]$$

Now take any field ϕ_{i_l} and decompose into $\phi_{i_l}^+$ and $\phi_{i_l}^-$. Commutator of ϕ_1^+ with the first term is simply zero while the commutator with the second term produces a contraction D_{1i_l}. In this way, the commutator part produces all contractions containing ϕ_1. Thus, the entire term produces all contractions with fields $\phi_1 \dots \phi_n$ and we showed that it is true for n fields, starting with the assumption that it is true for $n-1$ fields. By induction this is true for all $n \geq 1$ and hence proved.

Note that when we take the vacuum expectation of the time-ordered product, all the terms with a dangling normal ordering is killed, since the creation and annihilation operators act on the vacuum from the right direction so as to kill it. Only those term survive where all the fields have been contracted. For $n = 4$ only the blue terms survive which can be diagrammatically represented as

$$\langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\phi_3\phi_4\}\,|\Omega\rangle =$$

The diagrams represents the process of particle creation, propagation, and annihilation which takes place in spacetime. Each line denotes a propagator and the nodes are the starting or ending point of each process.

Lecture 31: Feynman diagrams

In the previous case, we discussed about Wick's theorem and time-ordered product of fields can be simply written as a sum of contractions of the fields. We will now consider the case of ϕ^4 interacting theory and try to find the n-point correlation functions of the interacting theory.

Recall from Eq. 43 that the correlation functions are basically the vacuum expectation of the time-ordered field products and the time-ordered product is known from Wick's theorem. Thus we already have a recipe to calculate this and we will use a perturbative approach for this.

We had considered the interaction of the form ϕ^4 and from this $\mathcal{H}_{\text{I}} = \frac{\lambda}{4!}\phi_{\text{I}}^4$ where everything is in the interaction picture. Let us consider the **numerator** for the two-point correlation function first which is of the form

$$\langle\Omega|\,\mathcal{T}\left\{\phi_1\phi_2 \exp\left(\frac{-i\lambda}{4!}\int \mathrm{d}^4z\ \phi^4(z)\right)\right\}|\Omega\rangle$$

We first Taylor expand the exponential

$$\exp\left(-\frac{i\lambda}{4!}\int \mathrm{d}^4z\ \phi^4(z)\right) = \mathbb{1} + \left(\frac{-i\lambda}{4!}\right)\int \mathrm{d}^4z\ \phi^4(z) + \left(\frac{-i\lambda}{4!}\right)^2 \int \mathrm{d}^4z\ \mathrm{d}^4w\ \phi^4(z)\phi^4(w) + \dots$$

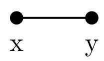
and then the numerator takes on a ghastly form

$$\langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\}\,|\Omega\rangle + \left(-\frac{i\lambda}{4!}\right)\int \mathrm{d}^4z\ \langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\phi^4(z)\}\,|\Omega\rangle + \frac{1}{2!}\left(-\frac{i\lambda}{4!}\right)^2\int \mathrm{d}^4z\mathrm{d}^4w\,\langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\phi^4(z)\phi^4(w)\}\,|\Omega\rangle + \dots$$

It is now possible to use Wick's theorem to characterise all these integrals, carefully considering the cumbersome contractions.

▷ Term of $\mathcal{O}(1)$:

The term is $\langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\}\,|\Omega\rangle$, which is simply $\triangle_{12}$ and is denoted diagrammatically as



▷ Term of $\mathcal{O}(\lambda)$:

We have to work a little bit here. The term is

$$\int \mathrm{d}^4z\ \langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\phi^4(z)\}\,|\Omega\rangle \equiv \int \mathrm{d}^4z\ \langle\Omega|\,\mathcal{T}\{\phi_1\phi_2\phi(z)\phi(z)\phi(z)\phi(z)\}\,|\Omega\rangle$$

We have to find all type of contractions amidst these six fields if we use the Wick's theorem. Note that any partial contractions will not matter, since we are taking vacuum expectations and uncontracted normal ordered products will kill it. Thus we need to consider only fully contracted fields. How many such configurations are there?

It turns out that we need to do some combinatorics for this. Indeed, all *good* math and physics ends in either linear algebra or combinatorics[1].

So if we take one field, then it can be contracted with the remaining 5, then these two fields are gone, remaining four are there. If we choose one, we can contract with remaining 3, and then the last two automatically gets contracted. So total $5 \times 3 = 15$ such fully contracted field configurations.

Due to symmetry conditions, we only have two distinct types of contractions which are,

$$\phi_1\phi_2\ \phi(z)\phi(z)\ \phi(z)\phi(z) \qquad \phi_1\phi_2\ \phi(z)\phi(z)\ \phi(z)\phi(z)$$

For the first case, we have $x-y$ contracting with each other and for the second case, we have $x-z$ contracting. These can be diagrammatically represented as:

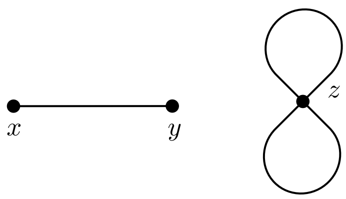


Figure 9: Type I contraction

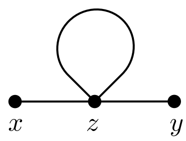


Figure 10: Type II contraction

There are three of the above Type I terms, since if x contracts with y then the first z can contract with three other guys.

Let us consider the second type of contraction where x contracts with z. There are 12 such terms since x can go with z in 4 ways and for each xz pair there are three ways y can go with z. Then the term of order $\mathcal{O}(\lambda)$ becomes

$$3 \times \left(\frac{-i\lambda}{4!}\right)\int \mathrm{d}^4z\ \triangle^2(x-y)\triangle^2(z-z) + 12 \times \left(\frac{-i\lambda}{4!}\right)\int \mathrm{d}^4z\ \triangle(x-z)\triangle(y-z)\triangle(z-z)$$

(For now just take $\triangle(z-z)$ to be some 'number' (even though it diverges) and let us not worry about this!)

Some observations

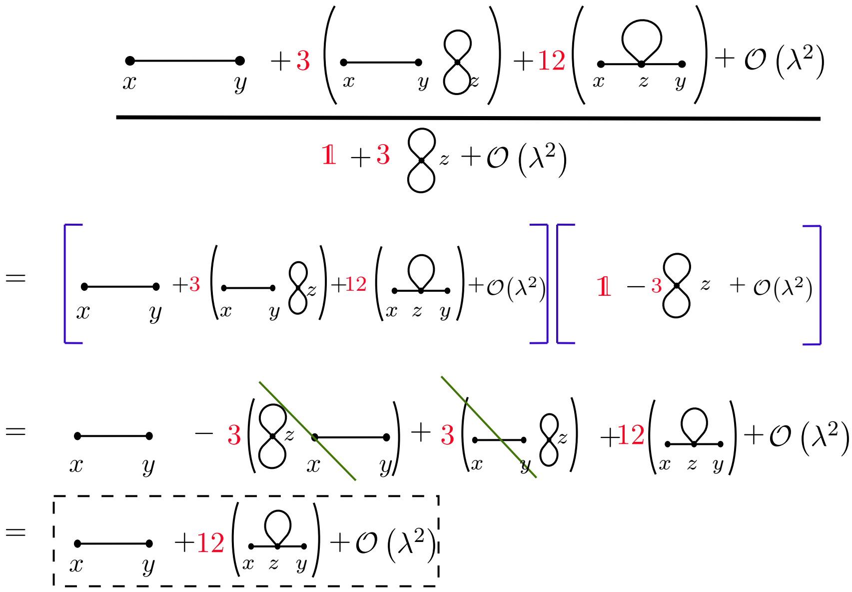
- ▷ The outermost nodes in the diagram (Eg. x and y) are the *external vertices*. These vertices are not integrated over.
- ▷ The inner nodes (Eg. z) are called the *internal vertices* and for each internal vertex, we have to associate an integral to them.

We had found the numerator upto order λ and now let us find the **denominator**.

$$\mathbb{1} + \left(-\frac{i\lambda}{4!}\right)\int \mathrm{d}^4z\ \langle\Omega|\,\mathcal{T}\{\phi^4(z)\}\,|\Omega\rangle + \frac{1}{2!}\left(-\frac{i\lambda}{4!}\right)^2\int \mathrm{d}^4z\mathrm{d}^4w\,\langle\Omega|\,\mathcal{T}\{\phi^4(z)\phi^4(w)\}\,|\Omega\rangle + \ldots$$

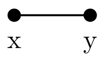
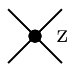
[1] which implies that if you don't find yourself doing these, you are not doing good physics

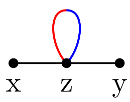
Note that the denominator does not have any free fields ϕ_1 and ϕ_2 and thus, the $\mathcal{O}(1)$ the term will be just $\mathbb{1}$ while the term of $\mathcal{O}(\lambda)$ will consist of only the contractions with z (3 terms). Thus symbolically (since the integrals will be just some number) we can write the two point correlator function as:



Note than, a cancellation was involved in the reduction such that the diagram with two disconnected components got cancelled. Even though this seems like a miracle, in fact, it's not!

Turns out, only those diagrams contribute where the components are connected to the external points. The contributions from all the disconnected subgraphs cancel. Before showing this, let us write a summary of what we need to do for systematic calculations:

- For each line  we associate the Feynman propagator $\triangle(x-y)$.
- For each internal vertex  we associate the term $(-i\lambda)\int d^4z$ (that is, we integrate over any internal vertices).
- For each external vertex, we associate the value 1.
- Finally we divide by the *symmetry factor*, which is like an *enigmatic term* but basically means the number of ways we can rearrange the propagators or vertices, without changing the diagram (Peskin says that we *almost surely* never need to calculate any diagram with symmetry factor greater than 2). For example, in the diagram

  we can interchange the blue and the red lines (propagators) which would keep the diagram intact. Hence there are two ways of arrangement and hence, the symmetry factor with which we divide the expression is 2.

  This is consistent with our explicit calculation, since we get a contribution of 12 for this diagram and then we divide by the 4! which is already there in the integral, giving us $^{12}/_{24} = {}^{1}/_{2}$.

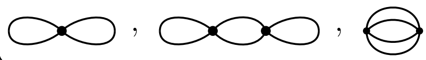
The above set of rules are called the *Feynman rules* in position-space (since we are considering the fields in the real space) and these are specific to the ϕ^4-theory. For other type of interactions, few changes are accordingly made.

Theorem 4 (Linked-Cluster Theorem):

$$\langle\widetilde{\Omega}|\,\mathcal{T}\{\phi_1\ldots\phi_n\}\,|\widetilde{\Omega}\rangle = \begin{pmatrix}\text{sum of all connected diagrams}\\ \text{subjected to } \textit{Feynman rules}\end{pmatrix}$$

Proof. A typical diagram consists of *connected* and *disconnected* diagrams, where connected diagrams are the ones where everything is connected to some external vertices. Disconnected diagrams are such that the diagrams consist of different components where at least one component is not connected to any external vertex. The components not connected to any external vertices are called the *vacuum bubbles.*

Let us consider the set of all vacuum bubbles represented by $\mathcal{V}$, then we have

$\mathcal{V} = \{$  , etc.$\}$

Suppose a diagram V (where V actually denotes the value of the integral represented by the diagram) has n_i pieces of $v_i \in \mathcal{V}$, in addition to the connected components (containing external vertices) whose value is, let say, c. Then the contribution of the diagram becomes:

$$c\cdot\prod_i \frac{1}{n_i!}(v_i)^{n_i}$$

where the $n_i!$ is taken since there are n_i copies of v_i, but these copies are indistinguishable. Then the numerator of the correlation function is obtained by summing over all the different connected components c

$$\text{numerator} \;=\; \sum_{\substack{\text{all possible}\\ \text{connected}\\ \text{pieces } c}}\;\sum_{\{n_i\}} c\cdot\prod_i \frac{1}{n_i!}(v_i)^{n_i} = \sum_{\substack{\text{all possible}\\ \text{connected}\\ \text{pieces } c}} c\cdot\prod_i\sum_{\{n_i\}}\frac{1}{n_i!}(v_i)^{n_i} = \left(\sum_{\substack{\text{all possible}\\ \text{connected}\\ \text{pieces } c}} c\right)\cdot\left(\exp\left(\sum_i v_i\right)\right)$$

where we have used the expansion for e^x and we finally obtain that the numerator is just the product of the connected contribution and the vacuum bubble contribution. The denominator has no connected contributions since the external fields do not occur in the denominator. Hence,

$$\text{denominator} \;=\; \exp\left(\sum_i v_i\right)$$

The exponential part thus cancels when we divide the numerator and the denominator, leaving only the contribution from the connected components.

NOTE: Higher order (>2 point) correlation functions may have apparent disconnected pieces in the diagram, like say, the diagrams for the four point correlation function in 30.2, where two of the diagrams are not connected. However, each disconnected piece is itself connected to some or the other externel vertices. When we talk about disconnected pieces (or vacuum bubbles), we talk about diagrams which contain a piece where that piece has no connection with any of the external vertex.

Lecture 32: Feynman diagrams in Momentum space

Recall that the Feynman propagator and its Fourier transform were defined as

$$D_{\text{F}}(x-y) = \int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\frac{i}{k^2-m^2+i\epsilon}e^{-ik\cdot(x-y)} \qquad \widetilde{D}_{\text{F}}(k) = \frac{i}{k^2-m^2+i\epsilon}$$

Then, if we convert the correlation function to the momentum space, we would have product of these propagators and also, we would introduce a delta function for any internal vertex. For example,

$$\int \mathrm{d}^4z \, \triangle(x-z)\triangle(y-z) \sim \int \mathrm{d}^4z \, \frac{\mathrm{d}^4p_1}{(2\pi)^4}\frac{\mathrm{d}^4p_2}{(2\pi)^4}\widetilde{D}_{\mathrm{F}}(p_1)\widetilde{D}_{\mathrm{F}}(p_2)e^{-ip_1\cdot x - ip_2\cdot y}e^{-i(p_1+p_2)\cdot z}$$
$$\sim \int \frac{\mathrm{d}^4p_1}{(2\pi)^4}\frac{\mathrm{d}^4p_2}{(2\pi)^4}\widetilde{D}_{\mathrm{F}}(p_1)\widetilde{D}_{\mathrm{F}}(p_2)\delta^{(4)}(p_1+p_2)e^{-ip_1\cdot x - ip_2\cdot y}$$

Note that the delta function in some sense, denotes *momentum conservation*. The Feynman rules for the correlation function in momentum space are as follows:

- For each internal line, we associate a momentum k and the propagator $\widetilde{D}_{\mathrm{F}}(k)$
- For each vertex, we write down the factor $(-i\lambda)(2\pi)^4\delta^{(4)}\left(\sum_i k_i\right)$ where $\sum k_i$ is the sum of all momenta flowing *into* the vertex that we are considering (*momentum conservation*).
- Divide by the symmetry factor and the other things that we did before.

The best thing would be to demonstrate through an example. Consider the diagram,

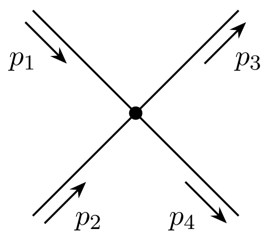


Figure 11: Diagram of order $\mathcal{O}(\lambda)$ for the four-point correlation function with interaction.

We can put the arrows of the momentum in any direction for now, the second rule will incorporate the sign automatically. The contribution of this diagram to the four-point correlation function, according to the above rules would be,

$$(-i\lambda)(2\pi)^4\left(\prod_{j=1}^{4}\frac{i}{p_j^2-m^2+i\epsilon}\right)\delta^{(4)}(p_1+p_2-p_3-p_4)$$

Now, consider a bit more involved diagram, containing a loop.

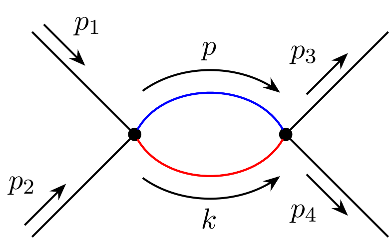


Figure 12: Diagram of order $\mathcal{O}(\lambda^2)$, containing a loop, for the four-point correlation function with interaction.

Let us first impose the momentum conservation, which will give us $\delta^{(4)}(p_1+p_2-p-k)$ and $\delta^{(4)}(p+k-p_3-p_4)$. Note that, these delta functions will be in product with each other. The argument of the second delta function becomes zero for $p = p_3+p_4-k$ and we can substitute this in the first delta function, leading to $\delta^{(4)}(p_1+p_2-p_3-p_4)$.

We see that, both p and k cannot be uniquely determined just from momentum conservation and hence, there is always a dangling momentum remaining, which we need to integrate out.

Also, we can interchange the red and blue lines, leaving the diagram unchanged, giving us a symmetry factor of 2. Thus, considering all the facts, the contribution of the diagram becomes:

$$\frac{1}{2}(-i\lambda)^2(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\epsilon}\frac{i}{(p_3+p_4-k)^2-m^2+i\epsilon}\left(\prod_{j=1}^{4}\frac{i}{p_j^2-m^2+i\epsilon}\right)$$

32.1. Scattering Matrix

We now focus on how QFT can be used to extract information from physical processes, particularly *scattering.* Consider a state $|i\rangle$ which is at $t\to-\infty$, describing particles, which evolves in time and goes to a final state $|f\rangle$ at $t\to+\infty$.

In the initial and final states, all the particles are out of each others' influence, however, in an intermediate time, these come together and *interact.* The amplitude for this process is given by:

$$\lim_{\tau\to\infty}\langle f|U(\tau,-\tau)|i\rangle\equiv\langle f|\mathcal{S}|i\rangle$$

where $\mathcal{S}$ is called the *scattering matrix* or the *S-matrix.* $\mathcal{S}$ is an unitary operator. In almost all cases, we will neglect the situation where no scattering has occurred, in which case we can define

$$\mathcal{S}=\mathbb{1}+i\mathcal{T}$$

and we will focus on the transfer matrix $\mathcal{T}$ henceforth. Till now we have been using the correlation function for all our interactions, however, it seems that the scattering amplitude is a more suitable quantity to measure in experimental observations.

There exists a thing called **Lehmann-Symanzik-Zimmermann (LSZ)** reduction formula which relates the correlation function with the scattering amplitude. The proof seems to be very complicated (for me), hence I am just going to mention the idea behind it.

Basically, for scattering amplitude, you are really interested in the actual *interaction* however, the correlation function also contains the free propagators.

What I mean is that, consider the diagram with the loop. For scattering processes, we generally know the initial and final momenta (momenta with which the particles are entering or going out), so we do not need to care about them. Thus, from the contribution, we remove the propagators of these 'known' momenta. This process is called *amputation*, where we remove the contribution from the external legs and this gives us the scattering matrix element.

There is always a delta function imposing momentum conservation which might become cumbersome to write everytime, hence we define the matrix $\mathcal{M}$ such that,

$$\langle\mathbf{p}_1\ldots\mathbf{p}_n|i\mathcal{T}|\mathbf{k}_1\ldots\mathbf{k}_m\rangle=(2\pi)^4\delta^{(4)}\left(\sum_i p_i-\sum_j k_j\right)(i\mathcal{M})$$

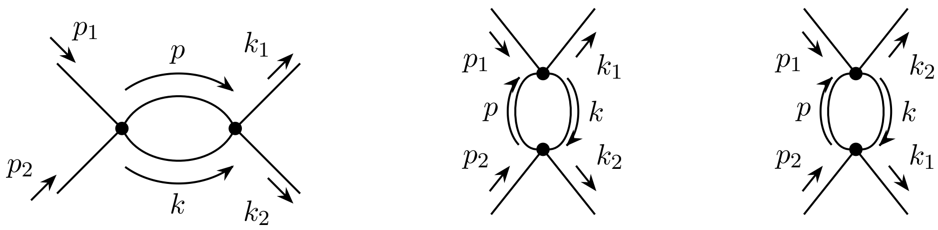
where p_i and k_j are the final and initial momenta. The matrix $\mathcal{M}$ will have contributions from the non-external propagators of all the diagrams that is possible for the process. In a gist, we have

$$\boxed{\mathcal{M}\equiv\text{ sum of all connected, amputated diagrams}}$$

For example, consider a process

$$p_1+p_2\longrightarrow k_1+k_2$$

Apart from the $\mathcal{O}(\lambda)$ diagram shown before, there are three order $\mathcal{O}(\lambda^2)$ diagrams which can represent the above scattering process and we have to consider all three of them:



- In the first diagram, momentum conservation gives $p_1 + p_2 = p + k = p_3 + p_4 \implies p = p_1 + p_2 - k$
- In the second diagram, momentum conservation gives $p_1 + p = k_1 + k \implies p = k_1 + k - p_1$
- In the first diagram, momentum conservation gives $p_1 + p = k_2 + k \implies p = k_2 + k - p_1$

The $\mathcal{O}(\lambda)$ process will contribute a factor of only $-i\lambda$ in the scattering matrix, since it has no loops (such a diagram is called a *tree-level diagram*) and hence no undetermined momenta to integrate out. Let us define the following quantity,

$$\mathcal{A}(p) \equiv \frac{(-i\lambda)^2}{2} \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(p-k)^2 - m^2 + i\epsilon}$$

Using $\mathcal{A}(p)$, we can write the matrix $\mathcal{M}$ upto $\mathcal{O}(\lambda^2)$ for the scattering process $p_1 + p_2 \longrightarrow k_1 + k_2$:

$$i\mathcal{M} = \underbrace{(-i\lambda)}_{\text{tree-level}} + \mathcal{A}(p_1 + p_2) + \mathcal{A}(p_1 - k_1) + \mathcal{A}(p_1 - k_2)$$

Lecture 33: Scalar Yukawa Theory

We have dealt with ϕ^4 interaction in the previous few sections and we saw some Feynman diagrams for the interaction. Now we will consider the interaction between a real scalar and a complex scalar field, which goes by the name of *Yukawa interaction* which is of the form:

$$\mathcal{H} = g\psi^\dagger \psi \phi$$

where ψ is our charged complex field and ϕ is the real scalar field. The full Lagrangian is given by:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 + (\partial_\mu \psi^\dagger)(\partial^\mu \psi) - M^2 \psi^\dagger \psi - g\psi^\dagger \psi \phi$$

where m is the mass associated with the ϕ field and M is the mass associated with the ψ field. The $1/2$ factor does not come with the complex scalar field, since a complex scalar field already has two degrees of freedom. And regarding the interaction, the simplest case is considered. $\psi^\dagger \psi$ gives a real, Lorentz invariant quantity and we multiply it with ϕ to get us an interaction between the two fields.

Having been familiarised with Feynman diagrams before, we will straightaway jump to calculate the amplitude of some scattering processes for this interaction. Before that we need to consider a few things:

- We have two kinds of field here, so we need a way to distinguish between them. We will denote the real scalar field by *dashed lines* while the complex scalar field by *solid lines.* Also, the solid lines will be *directed*, to denote the charge flow of the complex scalar field.

  The convention will be to have an *incoming* arrow for the initial state in ψ. For the initial state in $\psi^\dagger$, we will have an *outgoing* arrow. For the final state, the convention is reversed.

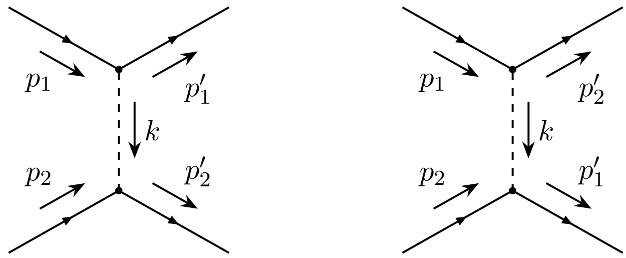
- Each interaction vertex will have *three* fields, since the interaction term has three fields. In ϕ^4 theory, we have tetravalent interaction vertices since the interaction had four terms in ϕ.

- The Feynman rules mostly remain same in this case. The propagator for real and complex field is the same, only the appropriate mass of the field needs to be taken in each case. Each vertex will have a factor of $(-ig)$.

33.1. Nucleon-Nucleon scattering

Consider the process $\underline{\psi\psi \longrightarrow \psi\psi}$ where two incoming ψ particles result in two outgoing ψ particles. Henceforth, we will call ψ particles are *nucleons* and ϕ particles as *mesons*. Hence this is a case of nucleon-nucleon scattering.

At order $\mathcal{O}(g^0)$ we have no scattering, both ψ comes and leaves as it is, without any interaction. Diagram of order $\mathcal{O}(g)$ with only one vertex is not possible for this process. Hence, the only meaningful diagram will start at order $\mathcal{O}(g^2)$. The simplest diagrams contributing to this process are,



Explicitly written with the momenta, this scattering process reads

$$\psi(p_1) + \psi(p_2) \longrightarrow \psi(p_1') + \psi(p_2')$$

Consider the first diagram. Momentum conservation leads to $k = p_1 - p_1' = p_2' - p_2$. Since there are two vertices, the contribution of the vertex will be $(-ig)^2$. We will consider only the propagator for k, since all the other momenta are determined. Then the scattering amplitude for the first diagram will be,

$$i\mathcal{M}_1 = (-ig)^2 \frac{i}{(p_1 - p_1')^2 - m^2 + i\epsilon}$$

where we have substituted $k = p_1 - p_1'$ and since k is associated with the ϕ field, we considered the mass m. Similarly, for the second diagram we have,

$$i\mathcal{M}_2 = (-ig)^2 \frac{i}{(p_1 - p_2')^2 - m^2 + i\epsilon}$$

The total scattering amplitude is the sum of these two terms,

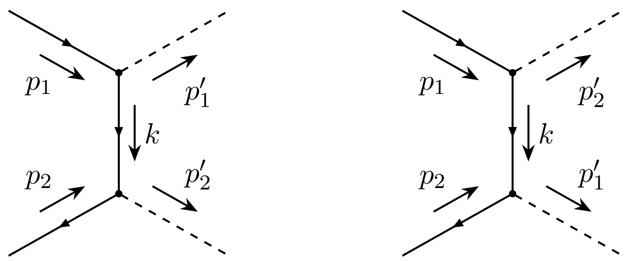
$$i\mathcal{M} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - m^2 + i\epsilon} + \frac{i}{(p_1 - p_2')^2 - m^2 + i\epsilon}\right]$$

These diagrams are tree-level diagrams and have no loops and hence no undetermined momentum to integrate over. The loops come in at $\mathcal{O}(g^4)$ and higher orders.

33.2. Nucleon to Meson Scattering

Let us consider another process,

$$\psi(p_1) + \psi^\dagger(p_2) \longrightarrow \phi(p_1') + \phi(p_2')$$



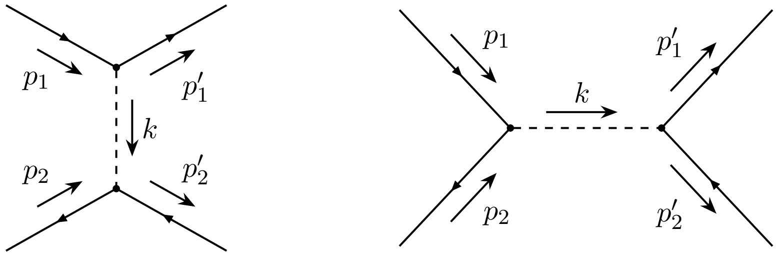
Observe the direction of the arrows denoting the charges. These have been drawn taking care of the *charge conservation* at each vertices. Since the internal line is now a *nucleon*, the amplitude is,

$$i\mathcal{M} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - M^2 + i\epsilon} + \frac{i}{(p_1 - p_2')^2 - M^2 + i\epsilon}\right]$$

33.3. Nucleon-Anti-Nucleon Scattering

Consider yet another process, which will have a bit different diagram than the previous two

$$\psi(p_1) + \psi^\dagger(p_2) \longrightarrow \psi(p_1') + \psi^\dagger(p_2')$$



The second diagram is a tad different from the other diagrams that we have seen. Momentum conservation in the second diagram gives $p_1 + p_2 = k$. Thus the amplitude for the process will be given by,

$$i\mathcal{M} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - m^2 + i\epsilon} + \frac{i}{(p_1 + p_2)^2 - m^2 + i\epsilon} \right]$$

33.4. Mandelstam variables

The quantities $(p_1 - p_1'), (p_1 - p_2')$ and $(p_1 + p_2)$ seems to appear in the propagator expression quite often. These are given some standard names, called *Mandelstam variables*, which are defined as the Lorentz invariant quantities,

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

These are also called s, t, u-channels. Firstly note that we have the total momentum conservation, $p_1 + p_2 = p_1' + p_2'$. Then,

$$\begin{aligned} s + t + u &= (p_1 + p_2)^2 + (p_1 - p_1')^2 + (p_1 - p_2')^2 \\ &= p_1^2 + p_2^2 + 2p_1 \cdot p_2 + p_1^2 + p_1'^2 - 2p_1 \cdot p_1' + p_1^2 + p_2'^2 - 2p_1 \cdot p_2' \\ &= 3p_1^2 + p_2'^2 + p_1'^2 + p_2^2 + 2p_1 \cdot (p_2 - p_1' - p_2') \\ &= 3p_1^2 + p_2'^2 + p_1'^2 + p_2^2 - 2p_1^2 \\ &= p_1^2 + p_2'^2 + p_1'^2 + p_2^2 \end{aligned}$$

where $p_i^2 \equiv (p_i)^\mu (p_i)_\mu$ and $p_i \cdot p_j \equiv (p_i)^\mu (p_j)_\mu$ are the Lorentz scalar product. Using the on-shell condition, $p_j^2 = m_j^2$ we have $s + t + u = \sum m_j^2$, which is a nice thing perhaps. Now, consider the *centre of mass/momentum* frame, where the total three-momentum is zero. Then,

$$\begin{aligned} \mathbf{p}_1 + \mathbf{p}_2 = 0 &\implies \mathbf{p}_i \equiv \mathbf{p}_1 = -\mathbf{p}_2 \\ \mathbf{p}_1' + \mathbf{p}_2' = 0 &\implies \mathbf{p}_f \equiv \mathbf{p}_1' = -\mathbf{p}_2' \end{aligned}$$

From the above conditions, the momentum four vectors become

$$p_1 = (E_1, \mathbf{p}_i), p_2 = (E_2, -\mathbf{p}_i), p_1' = (E_1', \mathbf{p}_f), p_2' = (E_2', -\mathbf{p}_f)$$

Then, $s = (p_1 + p_2)^2 \equiv (E_1 + E_2)^2 \equiv E_{\text{T}}^2$. s thus measures the total center of mass energy of the collision.

We can now find an expression for the energies in terms of the variable s. For that, we will require the on-shell condition, $p^2 = m^2$ and the above relation, which gives us:

$$E_1 + E_2 = \sqrt{s} \qquad E_1^2 - \mathbf{p}_i^2 = m_1^2 \qquad E_2^2 - \mathbf{p}_i^2 = m_2^2$$

Using the above, we get the following,

$$m_1^2 - m_2^2 = E_1^2 - E_2^2 = E_1^2 - (\sqrt{s} - E_1)^2 = \sqrt{s}(2E_1 - \sqrt{s}) \implies E_1 = \frac{s + m_1^2 - m_2^2}{2\sqrt{s}}$$

Similarly, we obtain the other energies in terms of the masses and s,

$$E_2 = \frac{s + m_2^2 - m_1^2}{2\sqrt{s}} \qquad E_1' = \frac{s + m_1'^2 - m_2'^2}{2\sqrt{s}} \qquad E_2' = \frac{s + m_2'^2 - m_1'^2}{2\sqrt{s}}$$

Let us now consider the variable $t = (p_1 - p_1')^2$ which turns out to be

$$\begin{aligned} t &= p_1^2 + p_1'^2 - 2p_1 \cdot p_1' \\ &= p_1^2 + p_1'^2 - 2((p_1)^0(p_1')^0 - |\mathbf{p}_1||\mathbf{p}_1'|\cos\theta) \\ &= p_1^2 + p_1'^2 - 2(E_1E_1' - |\mathbf{p}_1||\mathbf{p}_1'|\cos\theta) \end{aligned}$$

where θ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_1'$. Let us take the elastic limit, $m_1 = m_1'$ and $m_2 = m_2'$ which gives us $E_1 = E_1' = \mathfrak{E}_1$ and $E_2 = E_2' = \mathfrak{E}_2$ along with $|\mathbf{p}_i| = |\mathbf{p}_f| \equiv \mathfrak{p}$. Then the expression for t becomes,

$$t = \cancel{2\mathfrak{E}_1^2} - 2\mathfrak{p}^2 - \cancel{2\mathfrak{E}_1^2} + 2\mathfrak{p}^2\cos\theta = -2\mathfrak{p}^2(1 - \cos\theta) = -4\mathfrak{p}^2\sin^2\left(\frac{\theta}{2}\right) \leq 0$$

In a similar way, we can derive $u = -4\mathfrak{p}^2\cos^2\left(\frac{\theta}{2}\right) \leq 0$. Thus, t and u are measures of the momentum exchanged between particles. As the Mandelstam variables are scalars, these will retain these properties in all frames, so $s \geq 0$ and $t, u \leq 0$ in all frames.

The amplitudes of the scattering processes can be written entirely in terms of the Mandelstam variables.

- For the process $\psi\psi \longrightarrow \psi\psi$ we have $i\mathcal{M} = (-ig)^2\left[\frac{i}{t-m^2} + \frac{i}{u-m^2}\right]$
- For the process $\psi\psi^\dagger \longrightarrow \phi\phi$ we have $i\mathcal{M} = (-ig)^2\left[\frac{i}{t-M^2} + \frac{i}{u-M^2}\right]$
- For the process $\psi\psi^\dagger \longrightarrow \psi\psi^\dagger$ we have $i\mathcal{M} = (-ig)^2\left[\frac{i}{t-m^2} + \frac{i}{s-m^2+i\epsilon}\right]$

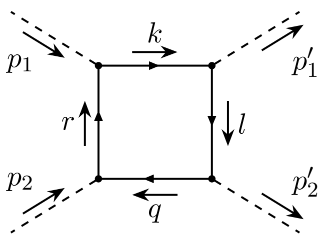
As t and u are always negative, the denominator in the amplitudes (except the red term) cannot be zero and hence we have safely removed the $+i\epsilon$ term. However, the red term with s channel can be zero, as $s \geq 0$. Take the COM frame for the s channel diagram. Since the particles are all nucleons, $s = E_{\mathrm{T}}^2 = 4(M^2 + \mathfrak{p}^2)$ and hence the denominator becomes $4(M^2 + \mathfrak{p}^2) - m^2$. If $m < 2M$ then $4M^2 - m^2 > 0$ which implies that the denominator is positive, so we can drop the $+i\epsilon$ term. If $m > 2M$ then the amplitude corresponding to the second diagram diverges at some value of the momentum and we have to keep $+i\epsilon$.

33.5. Meson Scattering

This will be another different process with a different diagram.

$$\phi(p_1) + \phi(p_2) \longrightarrow \phi(p_1') + \phi(p_2')$$

There are two incoming and two outgoing ϕ lines and we have to put two ψ lines to each of them. If we think about this, the simplest diagram for this process will contain a loop (which increases the complexity)!



Since there is a loop, all the momentum cannot be uniquely determined solely from momentum conservation. Let us keep l fixed and try to express k, r, q in terms of l. Momentum conservation at each vertex gives:

$$\triangleright\ \boxed{k = l + p_1'} \quad \triangleright\ l = q + p_2' \implies \boxed{q = l - p_2'} \quad \triangleright\ p_1 + r = k = l + p_1' \implies \boxed{r = l + p_1' - p_1}$$

The scattering amplitude will have the integral over the undetermined momentum l and the product of propagators containing l. Since there are four vertices, there will be a contribution of $(-ig)^4$

$$(-ig)^4 \int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{i}{[l^2 - M^2 + i\epsilon]} \frac{i}{[(l+p_1')^2 - M^2 + i\epsilon]} \frac{i}{[(l-p_2')^2 - M^2 + i\epsilon]} \frac{i}{[(l+p_1'-p_1)^2 - M^2 + i\epsilon]}$$

The integral seems to be complicated. I don't know how to do this. Even Tong has left the integral like this, so it's a good time for me to stop too!

There are two other diagrams contributing to the above process which we will see later. Another thing to say is that, these amplitudes can also be *painstakingly* found out using Dyson formula and Wick's theorem, however, Feynman rules makes it really easy and elegant. And that is why we all love him![1]

Lecture 33: Spinor Yukawa Theory

We have seen interactions only between scalar fields uptil now. We had also seen fermion fields and gauge fields earlier. Hence we will discuss what happens when fermions get coupled to scalar fields, which is the actual *Yukawa theory*. The interaction is of the form

$$\mathcal{H}_{\text{int}} = g\phi\overline{\Psi}\Psi$$

where Ψ is a Dirac spinor and $\overline{\Psi}$ is the Dirac adjoint. We needed to make a Lorentz invariant scalar and hence the simplest interaction, as seen before, had to be of this form. The entire Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\phi)(\partial_\mu\phi) - \frac{1}{2}m^2\phi^2 + \overline{\Psi}(i\not{\partial} - M)\Psi - g\phi\overline{\Psi}\Psi$$

[1]The other set of more exciting Feynman rules about life can be found in his book *Surely You're Joking, Mr. Feynman!*

Appendices

A. Few BS about Lorentz Group

We will discuss about Lorentz group here. The Lorentz group, denoted by $O(1,3)$, is defined to be the group of all transformations (Lorentz transformation or LT) that preserves the Minkowski metric, that is,

$$\Lambda^{\mathrm{T}}\eta\Lambda = \eta \iff \eta_{\mu\nu}\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta} \tag{44}$$

Note that, as long as we are not using the indices, we are talking about the *abstract* group element. However, when we consider the expression using index notation, we are essentially considering the matrix representation of the group element using the *standard basis* of the Minkowski space[1].

From the above definition, we can instantly see that for any $\Lambda \in O(1,3)$, $\det\Lambda = \pm 1$. For $\det\Lambda = +1$, we refer to them as *proper LT* (denoted by L_+) and for $\det\Lambda = -1$ (denote by L_-), we refer to them as *improper LT*.
In Eqn. 44, if we take $\alpha = \beta = 0$, we get:

$$1 = \eta_{\mu\nu}\Lambda^{\mu}{}_{0}\Lambda^{\nu}{}_{0} = \left(\Lambda^{0}{}_{0}\right)^2 - \sum_{i=1}^{3}\left(\Lambda^{i}{}_{0}\right)^2 \implies \left(\Lambda^{0}{}_{0}\right)^2 \geq 1$$

Thus, we can have $\Lambda^{0}{}_{0} \geq 1$ which we refer to as *orthochronous* (denoted by $L^{\uparrow}$) or $\Lambda^{0}{}_{0} \leq -1$ which we call *antichronous* (denoted by $L^{\downarrow}$).

From the above two classifications using the signs of the determinant and $\Lambda^{0}{}_{0}$, we can thus have four disconnected components of the Lorentz group:

- $L_+^{\uparrow}$: These are the proper orthochronous LT. The identity $\mathbb{1}$ is contained in this subgroup. It is denoted by by $\mathrm{SO}^+(1,3)$ and is continuously connected to the identity, that is, we can build up any finite Lorentz transformation by making many consecutive small Lorentz transformations. Sometimes, as an abuse of terminology, this subgroup is itself called the *Lorentz group*, since this contains the only physical transformations.
- $L_-^{\uparrow}$: These are written as a product of the *parity* and $L_+^{\uparrow}$ transformations.
- $L_+^{\downarrow}$: These contain both time and space inversion and hence are written as a product of PT and $L_+^{\uparrow}$ transformation.
- $L_-^{\downarrow}$: These contain only time inversion and hence are written as a product of *time reversal* and $L_+^{\uparrow}$ transformation.

Thus, the entire Lorentz group can be written symbolically:

$$L = L_+^{\uparrow} \bigcup PL_+^{\uparrow} \bigcup PTL_+^{\uparrow} \bigcup TL_+^{\uparrow}$$

Note that except the proper orthochronous LT, none of the other component can form subgroups since these do not contain the identity element.

A.1. Spacetime Representation

For the time being, we will consider only the proper, orthochronous Lorentz group, that is, $\mathrm{SO}^+(1,3)$ which is a Lie group. Since this is continuously connected to the identity, let us consider an infinitesimal transformation from the identity given by,

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

[1] The Mindkowski space is denoted by $\mathbb{R}^{1,3}$ which is equal to $\mathbb{R}^4$ along with the associated metric $\eta_{\mu\nu}$. $(1,3)$ refers to the metric signature, that is, 1 positive entry and 3 negative entries viz. $\mathrm{diag}(+1,-1,-1,-1)$

Substituting this in Eq. 44 and retaining terms upto the first oder in ω, we have the following:

$$\begin{aligned}\eta_{\alpha\beta} &= \eta_{\mu\nu}(\delta^\mu{}_\alpha + \omega^\mu{}_\alpha)(\delta^\nu{}_\beta + \omega^\nu{}_\beta)\\ &= \eta_{\alpha\beta} + \delta^\mu{}_\alpha \eta_{\mu\nu}\omega^\nu{}_\beta + \eta_{\mu\nu}\delta^\nu{}_\beta\omega^\mu{}_\alpha\\ &= \eta_{\alpha\beta} + \delta^\mu{}_\alpha\omega_{\mu\beta} + \delta^\nu{}_\beta\omega_{\nu\alpha}\\ &= \eta_{\alpha\beta} + \omega_{\alpha\beta} + \omega_{\beta\alpha}\end{aligned}$$

From this, we get that infinitesimal displacement ω must be anti-symmetric, that is,

$$\boxed{\omega_{\alpha\beta} = -\omega_{\beta\alpha}}$$

We know that antisymmetric matrices have six independent elements (since diagonals are zero and upper-diagonal has $(16-4)/2 = 6$ elements and lower diagonal are negative of the upper diagonal) and hence the basis of antisymmetric matrices consists six elements. Of these six, three corresponds to *rotations* while the other three corresponds to *boosts*

Let us consider this basis of the second rank anti-symmetric tensor $\{\mathcal{J}^{\rho\sigma}\}$ which forms the *generators* of the transformation. These are antisymmetric in the indices ρ, σ and hence there are six of them. These can be written as:

$$(\mathcal{J}^{\rho\sigma})^\mu{}_\nu := i(\eta^{\mu\rho}\delta^\sigma{}_\nu - \eta^{\mu\sigma}\delta^\rho{}_\nu)$$

As ω is antisymmetric, we can write it as a linear combination of the basis elements, that is,

$$\omega^\mu{}_\nu = -i\Omega_{\rho\sigma}(\mathcal{J}^{\rho\sigma})^\mu{}_\nu$$

where $\Omega_{\rho\sigma} \in \mathbb{R}$ (since $\mathcal{J}^{\mu\nu}$ are purely imaginary, $i\mathcal{J}^{\mu\nu}$ becomes real and hence Λ becomes real as it should be) since Lorentz group is a real Lie group. We can write the infinitesimal transformation in the following way:

$$\Lambda = \mathbb{1} - i\Omega_{\rho\sigma}\mathcal{J}^{\rho\sigma}$$

Any finite Lorentz transformation can be generated by exponentiating the generators, that is,

$$\Lambda_{\text{finite}} = \exp\left(-\frac{i}{2}\Omega_{\rho\sigma}\mathcal{J}^{\rho\sigma}\right)$$

Once we found the generators, any element of the Lie algebra of the Lorentz group will be a linear combination of them. It is a good time for us to calculate the Lie bracket since it is an essential characteristic of the Lie algebra. Since any element is a linear combination of the generators, it is sufficient to calculate the Lie bracket for the generators only, which we can do by the explicit form that we wrote earlier.

$$\begin{aligned}[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}]^\alpha{}_\beta &= (\mathcal{J}^{\mu\nu}\mathcal{J}^{\rho\sigma})^\alpha{}_\beta - (\mathcal{J}^{\rho\sigma}\mathcal{J}^{\mu\nu})^\alpha{}_\beta\\ &= (\mathcal{J}^{\mu\nu})^\alpha{}_\theta(\mathcal{J}^{\rho\sigma})^\theta{}_\beta - (\mathcal{J}^{\rho\sigma})^\alpha{}_\theta(\mathcal{J}^{\mu\nu})^\theta{}_\beta\\ &= (\eta^{\mu\alpha}\delta^\nu{}_\theta - \eta^{\nu\alpha}\delta^\mu{}_\theta)(\eta^{\rho\theta}\delta^\sigma{}_\beta - \eta^{\sigma\theta}\delta^\rho{}_\beta) - (\eta^{\rho\alpha}\delta^\sigma{}_\theta - \eta^{\sigma\alpha}\delta^\rho{}_\theta)(\eta^{\mu\theta}\delta^\nu{}_\beta - \eta^{\nu\theta}\delta^\mu{}_\beta)\\ &= \left\{\eta^{\mu\alpha}\delta^\nu{}_\theta\eta^{\rho\theta}\delta^\sigma{}_\beta - \eta^{\mu\alpha}\delta^\nu{}_\theta\eta^{\sigma\theta}\delta^\rho{}_\beta - \eta^{\nu\alpha}\delta^\mu{}_\theta\eta^{\rho\theta}\delta^\sigma{}_\beta + \eta^{\nu\alpha}\delta^\mu{}_\theta\eta^{\sigma\theta}\delta^\rho{}_\beta\right\}\\ &\quad - \left\{\eta^{\rho\alpha}\delta^\sigma{}_\theta\eta^{\mu\theta}\delta^\nu{}_\beta - \eta^{\rho\alpha}\delta^\sigma{}_\theta\eta^{\nu\theta}\delta^\mu{}_\beta - \eta^{\sigma\alpha}\delta^\rho{}_\theta\eta^{\mu\theta}\delta^\nu{}_\beta + \eta^{\sigma\alpha}\delta^\rho{}_\theta\eta^{\nu\theta}\delta^\mu{}_\beta\right\}\\ &= \left\{\eta^{\mu\alpha}\eta^{\rho\nu}\delta^\sigma{}_\beta - \eta^{\mu\alpha}\eta^{\sigma\nu}\delta^\rho{}_\beta - \eta^{\nu\alpha}\eta^{\rho\mu}\delta^\sigma{}_\beta + \eta^{\nu\alpha}\eta^{\sigma\mu}\delta^\rho{}_\beta\right\}\\ &\quad - \left\{\eta^{\rho\alpha}\eta^{\mu\sigma}\delta^\nu{}_\beta - \eta^{\rho\alpha}\eta^{\nu\sigma}\delta^\mu{}_\beta - \eta^{\sigma\alpha}\eta^{\mu\rho}\delta^\nu{}_\beta + \eta^{\sigma\alpha}\eta^{\nu\rho}\delta^\mu{}_\beta\right\}\text{complete proof!}\end{aligned}$$

From this, we find out the Lie bracket between the generators of the Lie Algebra of SO(1, 3):

$$[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}] = i(\eta^{\nu\rho}\mathcal{J}^{\mu\sigma} - \eta^{\mu\rho}\mathcal{J}^{\nu\sigma} - \eta^{\nu\sigma}\mathcal{J}^{\mu\rho} + \eta^{\mu\sigma}\mathcal{J}^{\nu\rho})$$

We had obtained the six generators for the Lie algebra and how let us rearrange these into two separate parts:

$$J^i = \frac{i}{2}\varepsilon^{ijk}\mathcal{J}^{jk} \qquad K^i = i\mathcal{J}^{i0}$$

In terms of these newly defined quantitites, the Lie bracket becomes check!:

$$\begin{aligned}\left[J^i, J^j\right] &= i\varepsilon^{ijk}J^k \\ \left[K^i, K^j\right] &= -i\varepsilon^{ijk}J^k \\ \left[J^i, K^j\right] &= i\varepsilon^{ijk}K^k\end{aligned}$$

This shows that J^i as defined above kinda denotes angular momentum since it satisfies the Lie Algebra of SU(2) while the other relation shows that K^i is a spatial vector. Now let us define two quantitites:

$$\theta^i = \frac{1}{2}\varepsilon^{ijk}\Omega_{jk} \qquad \eta^i = \Omega^{i0}$$

Using this we can write the generators as:

$$\frac{1}{2}\Omega_{\mu\nu}\mathcal{J}^{\mu\nu} = \Omega_{12}\mathcal{J}^{12} + \Omega_{13}\mathcal{J}^{13} + \Omega_{23}\mathcal{J}^{23} + \Omega_{i0}\mathcal{J}^{i0} = \boldsymbol{\theta}\cdot\mathbf{J} - \boldsymbol{\eta}\cdot\mathbf{K}$$

Thus any Lorentz transformation can then be written as:

$$\boxed{\Lambda = \exp\left(-i\,\boldsymbol{\theta}\cdot\mathbf{J} + i\,\boldsymbol{\eta}\cdot\mathbf{K}\right)}$$

B. Maxwell's Equations from Electromagnetic Tensor

Maxwell's electromagnetic tensor is defined as,

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{45}$$

where $A^\mu = (\Phi, \mathbf{A})$ is the four-potential. Using the four potential, we can write the electric and magnetic fields as

$$\mathbf{E} = -\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t} \qquad \mathbf{B} = \nabla\times\mathbf{A}$$

Let us now find some components of the electromagnetic tensor,

$$F^{0i} = \partial^0 A^i - \partial^i A^0 = \partial_0 A^i + \partial_i A^0 = -\left[-\frac{\mathrm{d}}{\mathrm{d}t}A^i - \partial_i A^0\right] = -E^i$$

We also have the expression of the magnetic field as,

$$\begin{aligned}B^i = \varepsilon^{ijk}\partial^j A^k = \frac{1}{2}\left[\varepsilon^{ijk}\partial^j A^k + \varepsilon^{ijk}\partial^j A^k\right] &= \frac{1}{2}\left[\varepsilon^{ijk}\partial^j A^k - \varepsilon^{ikj}\partial^j A^k\right] \\ &= \frac{1}{2}\left[\varepsilon^{ijk}\partial^j A^k - \varepsilon^{ijk}\partial^k A^j\right] \\ &= \frac{1}{2}\varepsilon^{ijk}\left(\partial^j A^k - \partial^k A^j\right) \\ &= \frac{1}{2}\varepsilon^{ijk}F^{jk}\end{aligned}$$

where in one step we introduced a minus by interchanging the Levi-Civita position and then in the next step, we renamed dummy indices $j \leftrightarrow k$. Thus, the explicit matrix representation of the electromagnetic tensor becomes:

$$[F^{\mu\nu}] = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & B^3 & -B^2 \\ E^2 & -B^3 & 0 & -B^1 \\ E^3 & B^2 & -B^1 & 0 \end{pmatrix}.$$