A Reluctant Experimentalist's QFT Companion: Rigorous Derivations for the Rest of Us

Jeongsoo Park
University of Minnesota – Twin Cities
Minneapolis, Minnesota 55455, USA

December 18, 2025

Preface

I began my journey into quantum field theory (QFT) at the age of 22, during my final year of college. Two years later—now, as a second—year PhD student in physics—I have only just begun to feel truly at ease with the subject. So yes—I found quantum field theory *very* hard. Not only did its counterintuitive concepts challenge me, but the unfamiliar mathematics (where $AB \neq BA$) presented even greater hurdles to comprehension.

This realization led me to create this study guide for QFT, with two primary goals in mind: (1) to force myself to study more rigorously, and (2) to provide detailed calculations wherever possible, allowing beginners to engage with QFT comfortably—perhaps even from the comfort of their bed—building familiarity with both the ideas and the computations without second-guessing every step.

This guide is based on *Elementary Particle Physics Vol. 1*, written by Professor Yorikiyo Nagashima, who generously granted me permission to share it publicly in order to promote the accessibility of physics to a wider audience. If you seek deeper insights into quantum field theory, I highly recommend reading the original book as well.

I must also thank—and, frankly, apologize to—my PhD advisor, Professor Xianghan Xu, who has been endlessly patient with my slow understanding and greatly helpful in advancing my research, for indulging this little guilty hobby of mine. It is a profound privilege to have him as my advisor, guiding me through the twists and turns of both theory and experiment.

Compiling this guide has consumed a decent chunk of time that I could have spent in the lab fabricating a couple more feed rods instead. Maximizing my lab hours by shelving this project might even boost my paper output during my

PhD, improving my chances of securing a green card down the line. Nonetheless, I suspect my PhD years may be my last opportunity to devote myself so fully to studying QFT with such unbridled enthusiasm.

In total, I plan to produce three study guides: one on QFT, another on electroweak theory, and a third on quantum chromodynamics. This document represents the first one-third of the inaugural guide—there's still a long road ahead. But in case it benefits even one person on Earth, I've chosen to share it now and commit to timely updates as I progress.

Please keep in mind that this study guide was written by me, just another student striving to grasp these concepts while holding onto my sanity (and life). Approach it critically if you choose to read it—I would greatly appreciate any flaws you spot, so I can correct them and share your feedback. That said, I promise to pour the best of myself into making it as informative and useful as possible.

Acknowledgements

This study guide is inspired by the exceptional *Elementary Particle Physics Vol. 1*, authored by Professor Yorikiyo Nagashima. I extend my sincere gratitude to him for graciously permitting the free public distribution of this work, in the spirit of broadening access to quantum field theory. For deeper perspectives and foundational insights, I strongly encourage readers to explore the original book alongside this guide.

Contents

1 Motivation of Dirac equation and its conditions

Quantum mechanics tells you that a particle's dynamics can be described by the Schrodinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\frac{\vec{p}^2}{2m} + V(\vec{x})\right)\psi \tag{1}$$

where $p_i = -i\hbar\frac{\partial}{\partial x^i}$.

One obvious problem with Eq.(1) is that the time part and the spatial part are not treated equally. That is, Eq.(1) has the first derivative with respect to time but the second derivatives with respect to space. However, special relativity puts time and space on equal footing. So we need a relativistic version of Eq.(1). Before proceeding any further, we adopt natural units:

$$c = \hbar = 1 \tag{2}$$

1.1 Klein-Gordon Equation

According to special relativity, a particle's magnitude of 4-momentum $p^\mu = (E, \vec{p})$ is conserved and is equal to the squared mass of the particle.

$$E^2 - |\vec{p}|^2 = m^2 \tag{3}$$

In the context of special relativity, E is just a scalar, $\vec{p}$ is a vector (therefore $|\vec{p}|^2$ is a scalar), and m is a scalar. However, let us make the following quantum mechanical substitution (just as we did in non-relativistic quantum mechanics):

$$E \to i\frac{\partial}{\partial t} \qquad\qquad p_i \to -i\frac{\partial}{\partial x^i} \tag{4}$$

If we blindly perform the substitution (4) into Eq.(3), we have:

$$-\frac{\partial^2}{\partial t^2} + \nabla^2 = m^2 \tag{5}$$

where

$$\nabla^2 = \sum_{i=1}^{n} \frac{\partial^2}{\partial x^{i^2}}$$

Eq.(5) does not make sense in an operator-wise manner. So we let the operators act on some wave function $\varphi(t, \vec{r})$. Then Eq.(5) becomes:

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\varphi(t, \vec{x}) = m^2\varphi(t, \vec{x}) \tag{6}$$

Eq.(6) is called the Klein-Gordon equation. Let's just take a simple look at Eq.(6). Well, it has second derivatives in both time and space. That sounds like an improvement. Let us try a plane wave solution, which is written as

$$\varphi(t,\vec{x}) = Ae^{-i(Et-\vec{p}\cdot\vec{x})} \tag{7}$$

where A is a normalization constant. By plugging Eq.(7) into Eq.(6),

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right) Ae^{-i(Et-\vec{p}\cdot\vec{x})} = m^2 Ae^{-i(Et-\vec{p}\cdot\vec{x})}$$

$$(E^2 - |\vec{p}|^2) Ae^{-i(Et-\vec{p}\cdot\vec{x})} = m^2 Ae^{-i(Et-\vec{p}\cdot\vec{x})}$$

Consequently, the equation above becomes

$$E^2 - |\vec{p}|^2 = m^2 \tag{8}$$

This is the energy-momentum relation from special relativity, which is what we want. Thus Eq.(6) admits the plane wave solution (Eq.(7)). The general solution of the Klein-Gordon equation can be expressed as a superposition of Eq.(7):

$$\Phi(t,\vec{x}) = \int_{-\infty}^{\infty} Ae^{-i(Et-\vec{p}\cdot\vec{x})} d^3p \tag{9}$$

(The reason why we do not integrate with respect to dE is that E is not an independent variable due to the energy-momentum relation $E^2 - |\vec{p}|^2 = m^2$.) So far, this is all mathematical discussion of the solution of the Klein-Gordon equation. Then what exactly *is* Eq.(7)? Well, just like we did in quantum mechanics, the first guess is: Eq.(7) is a wave function which satisfies the Klein-Gordon equation, and its probability density of the associated particle's position is given by some expression in terms of $\varphi(t,\vec{r})$. But is it correct? Let's calculate the probability density ρ.

To calculate ρ, we have to define how to calculate it. ρ is probability density which, we know, is locally conserved (just like charge density is locally conserved). Thus ρ is defined such that the continuity equation is satisfied.

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\vec{j} = 0 \tag{10}$$

Hence, we require that ρ and $\vec{j}$ are expressed in terms of $\varphi(x)$ and its derivatives to satisfy Eq.(10).

1.2 Problems of the Klein-Gordon equation

If we *were* dealing with the Schrodinger equation, ρ and $\vec{j}$ can be obtained by the following procedures. Start with the Schrodinger equation and its complex conjugate:

$$i\frac{\partial\psi}{\partial t} + \left(\frac{1}{2m}\nabla^2 - V(\vec{r})\right)\psi = 0 \tag{11}$$

$$-i\frac{\partial\psi^*}{\partial t} + \left(\frac{1}{2m}\nabla^2 - V(\vec{r})\right)\psi^* = 0 \tag{12}$$

Multiply $-i\psi^*$ to Eq.(11) and $-i\psi$ to Eq.(12) both from the left.

$$\psi^*\frac{\partial\psi}{\partial t} + \psi^*\left(\frac{-i}{2m}\nabla^2 + iV(\vec{r})\right)\psi = 0 \tag{13}$$

$$-\psi\frac{\partial\psi^*}{\partial t} + \psi\left(\frac{-i}{2m}\nabla^2 + iV(\vec{r})\right)\psi^* = 0 \tag{14}$$

Subtract Eq.(14) from Eq.(13).

$$\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t} - \frac{i}{2m}\psi^*\nabla^2\psi + \frac{i}{2m}\psi\nabla^2\psi^* = 0$$

We can combine the first two terms and the last two terms into the following:

$$\frac{\partial}{\partial t}(\psi^*\psi) - \frac{i}{2m}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*) = 0 \tag{15}$$

By comparing Eq.(10) and Eq.(15), we find:

$$\rho = \psi^*\psi \qquad\qquad \vec{j} = \frac{i}{2m}(\psi^*\nabla\psi - \psi\nabla\psi^*) \tag{16}$$

In this way, we did not have to assume that the probability density is given by $\psi^*\psi$. Instead, we proposed that the probability density obeys the continuity equation and found an expression for ρ and $\vec{j}$. All good, except we are dealing with the Klein-Gordon equation, rather than the Schrodinger equation. But we can follow a similar procedure.

1.2.1 Probability Density of the Klein-Gordon Equation

Let us start with the Klein-Gordon equation and its complex conjugate.

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\varphi = m^2\varphi \tag{17}$$

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\varphi^* = m^2\varphi^* \tag{18}$$

Multiply $-i\varphi^*$ to Eq.(17) and $-i\varphi$ to Eq.(18) from the left.

$$-i\varphi^*\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\varphi = -im^2\varphi^*\varphi \tag{19}$$

$$-i\varphi\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\varphi^* = -im^2\varphi\varphi^* \tag{20}$$

Subtract Eq.(20) from Eq.(19).

$$i\varphi^*\frac{\partial^2}{\partial t^2}\varphi - i\varphi\frac{\partial^2}{\partial t^2}\varphi^* - i\varphi^*\nabla^2\varphi + i\varphi\nabla^2\varphi^* = 0$$

In the equation above, we can combine the first two terms and the last two terms into the following:

$$\frac{\partial}{\partial t}\left[i\left(\varphi^*\frac{\partial}{\partial t}\varphi - \varphi\frac{\partial}{\partial t}\varphi^*\right)\right] + \nabla\cdot\left[i\left(\varphi\nabla\varphi^* - \varphi^*\nabla\varphi\right)\right] = 0 \tag{21}$$

By comparing Eq.(10) and Eq.(21), we find ρ and $\vec{j}$ as follows:

$$\rho = i\left(\varphi^*\frac{\partial}{\partial t}\varphi - \varphi\frac{\partial}{\partial t}\varphi^*\right) \qquad\qquad \vec{j} = i\left(\varphi\nabla\varphi^* - \varphi^*\nabla\varphi\right) \tag{22}$$

Eq.(22) is the Klein-Gordon version of Eq.(16). Let us plug the plane-wave solution (Eq.(7)) into Eq.(22) in order to obtain ρ. Then ρ is given by the following:

$$\rho = 2E|A|^2 \tag{23}$$

Looks all good, isn't it? However, relativistic energy of a particle is given by the energy-momentum relation, which is $E^2 = p^2 + m^2$, **not** $E = \sqrt{p^2+m^2}$. Therefore, energy can be either positive or, technically, negative: $E = \pm\sqrt{p^2+m^2}$. So far, it seems that there are two problems with the Klein-Gordon equation. (1) If $E < 0$, a particle can emit an infinite amount of energy in order to keep lowering its energy state. (2) If $E < 0$, then $\rho < 0$, which means that the probability density can be negative.

1.3 Dirac Equation

Let us be clear on why we had a negative probability density. From Eq.(23), the negative probability density comes from negative energy. This negative energy comes from the negative solution of $E^2 = \vec{p}^2 + m^2$ ($E = -\sqrt{\vec{p}^2+m^2}$). Note that E^2 *becomes* $-\frac{\partial^2}{\partial t^2}$ due to the quantum mechanical substitution (Eq.(4)). Thus we can say that the negative energy originates from the second derivative with respect to time! Thus our motivation is to come up with an equation which has the first derivative with respect to time.

1.3.1 Dirac Equation

Nevertheless, we have to start somewhere. So let's start with the Schrodinger equation:

$$i\frac{\partial}{\partial t}\psi = H\psi \tag{24}$$

Now, instead of $H = \frac{|\vec{p}|^2}{2m} + V(\vec{x})$, we let H be a linear combination of momentum and mass (so that time and space have equal order in derivatives).

$$\begin{aligned} H &= \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m \\ &= -i\alpha_1\frac{\partial}{\partial x^1} - i\alpha_2\frac{\partial}{\partial x^2} - i\alpha_3\frac{\partial}{\partial x^3} + \beta m \end{aligned} \tag{25}$$

By plugging Eq.(25) into Eq.(24), we have

$$\begin{aligned} i\frac{\partial}{\partial t}\psi &= (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)\psi \\ &= \left(-i\alpha_1\frac{\partial}{\partial x^1} - i\alpha_2\frac{\partial}{\partial x^2} - i\alpha_3\frac{\partial}{\partial x^3} + \beta m\right)\psi \end{aligned} \tag{26}$$

Before proceeding any further, let's take a look at Eq.(26). The first thing to note is that the equation has the first time derivative, which, we hope, eliminates the negative probability density that we had in the Klein-Gordon equation. The second thing is that the equation has the same order with respect to time and space (which is 1st-order). This agrees with special relativity, which treats space and time on equal footing. The mass term, βm, comes from the rest energy of a particle, which is mc^2. Eq.(26) is called the **Dirac equation**.

1.3.2 Dirac Conditions

We have not discussed what α_i $(i = 1, 2, 3)$ and β are. Are they specific numbers or some matrices? Let us investigate this. First, we require $E^2 = |\vec{p}|^2 + m^2$.

$$E^2 = |\vec{p}|^2 + m^2$$

By multiplying ψ from the right on both sides,

$$E^2\psi = (|\vec{p}|^2 + m^2)\psi$$

Keeping in mind the quantum mechanical substitution (Eq.(4)) and Eq.(26),

$$-\frac{\partial^2}{\partial t^2}\psi = (|\vec{p}|^2 + m^2)\psi = (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)^2\,\psi \tag{27}$$

From the second equality from Eq.(27), we can set (operator-wise):

$$\begin{aligned} p_1^2 + p_2^2 + p_3^2 + m^2 &= (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)^2 \\ &= \alpha_1^2 p_1^2 + \alpha_2^2 p_2^2 + \alpha_3^2 p_3^2 \\ &\quad + \alpha_1\alpha_2 p_1 p_2 + \alpha_1\alpha_3 p_1 p_3 + \alpha_2\alpha_1 p_2 p_1 + \alpha_3\alpha_1 p_3 p_1 + \alpha_2\alpha_3 p_2 p_3 + \alpha_2\alpha_3 p_2 p_3 \\ &\quad + (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3)\beta m + \beta m(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3) \\ &\quad + \beta^2 m^2 \\ &= \alpha_1^2 p_1^2 + \alpha_2^2 p_2^2 + \alpha_3^2 p_3^2 \\ &\quad + \{\alpha_1, \alpha_2\} p_1 p_2 + \{\alpha_1, \alpha_3\} p_1 p_3 + \{\alpha_2, \alpha_3\} p_2 p_3 \\ &\quad + \{\alpha_1, \beta\} p_1 m + \{\alpha_2, \beta\} p_2 m + \{\alpha_3, \beta\} p_3 m \\ &\quad + \beta^2 m^2 \end{aligned} \tag{28}$$

Note that p_i is basically a partial derivative operator, so p_i and p_j commute ($p_ip_j = p_jp_i$). By comparing LHS and RHS of Eq.(28),

$$\begin{aligned} &\text{(i)} \qquad \alpha_i^2 = \beta^2 = I \\ &\text{(ii)} \qquad \{\alpha_i, \alpha_j\} = 0 \qquad \text{if } i \neq j \\ &\text{(iii)} \qquad \{\alpha_i, \beta\} = 0 \end{aligned} \tag{29}$$

Eq.(29) represents the conditions that α_i and β must satisfy and is called the **Dirac conditions**. Let's assume that α_i and β are complex numbers. Then, from Eq.(29(iii)),

$$\alpha_i\beta + \beta\alpha_i = 2\alpha_i\beta = 0 \tag{30}$$

because two complex numbers always commute. Thus, by Eq.(30),

$$\alpha_i = 0 \qquad \text{or} \qquad \beta = 0 \tag{31}$$

However, Eq.(31) is contradicted by Eq.(29(i)). Thus, α_i and β cannot be complex numbers. So it is reasonable to assume that they are matrices. With Eq.(29), we can derive two properties of α_i and β.

(1) *Eigenvalues of α_i and β are ± 1.*

pf) $\alpha_i\psi = \lambda\psi \rightarrow \alpha_i^2\psi = \psi = \lambda^2\psi \qquad$ Thus, $\lambda^2 = 1 \rightarrow \lambda = \pm 1$

where we used $\alpha_i^2 = 1$ (Eq.(29.(i))).

(2) *α_i and β are traceless.*

pf) $\mathrm{Tr}[\alpha_i] = \mathrm{Tr}[\alpha_i\beta^2] = \mathrm{Tr}[\beta\alpha_i\beta] = -\mathrm{Tr}[\alpha_i\beta^2] = -\mathrm{Tr}[\alpha_i] \rightarrow \mathrm{Tr}[\alpha_i] = 0$

where we used $\beta^2 = I$ on the first equality, $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]$ on the second equality ($A = \alpha_i\beta$ and $B = \beta$), $\{\alpha_i, \beta\} = 0$ on the third equality, and, again, $\beta^2 = I$ on the fourth equality. Therefore, due to the properties (1) and (2), α_i and β must be some matrices of even dimension. (Try out for β!)

In fact, we can guess what α_i are from property (1). Can you remember some traceless matrices with eigenvalues of ± 1? We do – the Pauli matrices! The eigenvalues of the Pauli matrices are ± 1, which are eigenvalues of spin-up or spin-down state (omitting $\hbar/2$) along x, y, or z – axis. Here are the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{32}$$

Indeed, σ_x, σ_y, σ_z are traceless. **Note that $\alpha_i = -\sigma_i$ also works.** What about β? We will define β as a 4×4 matrix as well as σ_i after discussing the following example.

1.4 Weyl Particle

Starting with the Dirac equation (Eq.(26)), the Weyl particle is defined when $m = 0$.

$$\begin{aligned} i\frac{\partial}{\partial t}\psi &= (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3)\psi \\ &= \left(-i\alpha_1\frac{\partial}{\partial x^1} - i\alpha_2\frac{\partial}{\partial x^2} - i\alpha_3\frac{\partial}{\partial x^3}\right)\psi \\ &= \pm\vec{\sigma}\cdot(-i\nabla)\psi = \pm\vec{\sigma}\cdot\vec{p}\psi \end{aligned} \tag{33}$$

Here the Hamiltonian operator is $H = \pm\vec{\sigma}\cdot\vec{p}$. Let us assume that ψ is an eigenstate of the Hamiltonian operator. Then we have the following:

$$E\psi = \pm\vec{\sigma}\cdot\vec{p}\psi \tag{34}$$

So there are two cases depending on the sign in front of the Hamiltonian operator. So we have either

$$\text{(i)} \qquad E\psi = -\vec{\sigma}\cdot\vec{p}\psi \qquad \text{or} \qquad \text{(ii)} \qquad E\psi = \vec{\sigma}\cdot\vec{p}\psi \tag{35}$$

Let us write the solution of Eq.(35(i)) and Eq.(35(ii)) to be ϕ_L and ϕ_R, respectively. Thus we have:

$$\text{(i)} \qquad E\phi_L = -\vec{\sigma}\cdot\vec{p}\phi_L \qquad \text{or} \qquad \text{(ii)} \qquad E\phi_R = \vec{\sigma}\cdot\vec{p}\phi_R \tag{36}$$

By rewriting $\vec{p}$ as $|\vec{p}|\hat{p}$ and dividing both sides by E, Eq.(36) is:

$$\text{(i)} \qquad \phi_L = -\frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_L \qquad \text{or} \qquad \text{(ii)} \qquad \phi_R = \frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_R \tag{37}$$

Notice that $\frac{|\vec{p}|}{E}$ is a scalar. Thus, ϕ_L and ϕ_R are eigenstates of the operator $\vec{\sigma}\cdot\hat{p}$, which is called helicity operator. The helicity operator means spin projected onto the direction of momentum. Let us calculate the explicit form of the helicity operator and find its eigenvalues and eigenstates.

An arbitrary unit vector in spherical coordinate system is given by: $\hat{p} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z}$. Then the helicity operator $h = \vec{\sigma}\cdot\hat{p}$ is given by:

$$\begin{aligned} h = \vec{\sigma}\cdot\hat{p} &= \sigma_x\sin\theta\cos\phi + \sigma_y\sin\theta\sin\phi + \sigma_z\cos\theta \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sin\theta\cos\phi + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\sin\theta\sin\phi + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\cos\theta \\ &= \begin{pmatrix} \cos\theta & \sin\theta\cos\phi - i\sin\theta\sin\phi \\ \sin\theta\cos\phi + i\sin\theta\sin\phi & -\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} \end{aligned} \tag{38}$$

By inspection, the eigenvalue of h is given by ± 1. It is because a spin-1/2 particle can have either spin up or down along *any* direction $\hat{p}$. But let's confirm explicitly to see if this is the case.

$$\begin{aligned} &\begin{vmatrix} \cos\theta - \lambda & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta - \lambda \end{vmatrix} = (\cos\theta - \lambda)(-\cos\theta - \lambda) - \sin^2\theta = 0 \\ &\to \lambda^2 = 1 \to \lambda = \pm 1 \end{aligned} \tag{39}$$

If $\lambda = \pm 1$, the eigenstate $\chi_\pm$ is:

$$\begin{aligned} \lambda = +1 \to \chi_+ &= \begin{pmatrix} \cos\frac{\theta}{2} e^{-i\phi/2} \\ \sin\frac{\theta}{2} e^{i\phi/2} \end{pmatrix} \\ \lambda = -1 \to \chi_- &= \begin{pmatrix} -\sin\frac{\theta}{2} e^{-i\phi/2} \\ \cos\frac{\theta}{2} e^{i\phi/2} \end{pmatrix} \end{aligned} \tag{40}$$

So far, so good. let's have our attention on Eq.(37):

$$\text{(i)} \qquad \phi_L = -\frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_L \qquad \text{or} \qquad \text{(ii)} \qquad \phi_R = \frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_R$$

Since $m = 0$, $E^2 = |\vec{p}|^2 + m^2 = |\vec{p}|^2$. Therefore, $E = \pm|\vec{p}|$, which implies $\frac{|\vec{p}|}{E} = \pm 1$. Suppose we start with case(i) above:

$$\text{(i)} \qquad \phi_L = -\frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_L = \begin{cases} -\vec{\sigma}\cdot\hat{p}\phi_L & \text{if } E > 0 \\ +\vec{\sigma}\cdot\hat{p}\phi_L & \text{if } E < 0 \end{cases}$$

So if $E > 0$, ϕ_L is an eigenstate of the helicity operator and its eigenvalue is -1. Likewise, if $E < 0$, ϕ_L is an eigenstate of the helicity operator and its eigenvalue is $+1$. Combined with Eq.(40), this implies:

$$\text{(i)} \qquad -\vec{\sigma} \to \phi_L = \begin{cases} \chi_- & \text{if } E > 0 \to \text{Particle!} \\ \chi_+ & \text{if } E < 0 \to \text{Antiparticle!} \end{cases} \tag{41}$$

For now, let's say a particle with positive energy called a *particle* and with negative energy called an *antiparticle*. Let's do the similar analysis on case(ii).

$$\text{(ii)} \qquad \phi_R = \frac{|\vec{p}|}{E}\vec{\sigma}\cdot\hat{p}\phi_R = \begin{cases} +\vec{\sigma}\cdot\hat{p}\phi_R & \text{if } E > 0 \\ -\vec{\sigma}\cdot\hat{p}\phi_R & \text{if } E < 0 \end{cases}$$

If $E > 0$, ϕ_R is an eigenstate of the helicity operator and its eigenvalue is $+1$. Likewise, if $E < 0$, ϕ_L is an eigenstate of the helicity operator and its eigenvalue is -1. Again, combined with Eq.(40), this implies:

$$\text{(ii)} \qquad +\vec{\sigma} \to \phi_R = \begin{cases} \chi_+ & \text{if } E > 0 \to \text{Particle!} \\ \chi_- & \text{if } E < 0 \to \text{Antiparticle!} \end{cases} \tag{42}$$

Let us recap what was going on. The difference between ϕ_L and ϕ_R was whether we choose $-\sigma_i$ or $+\sigma_i$ for the matrices α_i which satisfy Eq.(29). If we choose $-\sigma_i$, then we have ϕ_L. Likewise, if we choose $+\sigma_i$, then we have ϕ_R. Whether ϕ_L or ϕ_R can be either χ_- or χ_+ depends on if the particle has positive or negative energy. Since σ_i is a 2×2 matrix, ϕ_L and ϕ_R must be a vector with two components. Is there any way to combine ϕ_L and ϕ_R with one equation? There is! Let's introduce a four-component vector ψ and 4×4 matrices α_i, and, finally, β.

$$\psi = \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} \qquad \alpha_i = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \qquad \beta = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \tag{43}$$

First of all, the extension (43) satisfies the Dirac conditions (Eq.(29)):

$$\begin{array}{lll} \text{(i)} & \alpha_i^2 = \beta^2 = I & \\ \text{(ii)} & \{\alpha_i, \alpha_j\} = 0 & \text{if } i \neq j \\ \text{(iii)} & \{\alpha_i, \beta\} = 0 & \end{array}$$

$$\alpha_i^2 = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_i^2 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix} = I$$

$$\beta^2 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix} = I$$

$$\begin{aligned} \{\alpha_i, \alpha_j\} &= \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} -\sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix} + \begin{pmatrix} -\sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix} \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \\ &= \begin{pmatrix} \{\sigma_i, \sigma_j\} & 0 \\ 0 & \{\sigma_i, \sigma_j\} \end{pmatrix} = 0 \qquad \text{if } i \neq j \end{aligned}$$

$$\begin{aligned} \{\alpha_i, \beta\} &= \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} = 0 \end{aligned}$$

Let's explicitly confirm if the extension (43) really reproduces Eq.(37).

$$E\psi = \begin{pmatrix} -\vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{p}\psi \tag{44}$$

where

$$\begin{pmatrix} -\vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{p} = \begin{pmatrix} -\sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix} p_1 + \begin{pmatrix} -\sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix} p_2 + \begin{pmatrix} -\sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} p_3$$

Then Eq.(44) becomes:

$$
\begin{aligned}
E\begin{pmatrix}\phi_L\\ \phi_R\end{pmatrix} &= \left[\begin{pmatrix}-\sigma_1 & 0\\ 0 & \sigma_1\end{pmatrix}p_1 + \begin{pmatrix}-\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}p_2 + \begin{pmatrix}-\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}p_3\right]\psi \\
&= \begin{pmatrix}-\vec{\sigma}\cdot\vec{p} & 0\\ 0 & \vec{\sigma}\cdot\vec{p}\end{pmatrix}\begin{pmatrix}\phi_L\\ \phi_R\end{pmatrix}
\end{aligned}
\tag{45}
$$

Indeed, the extension combines the two separate cases in Eq.(37).

1.5 Did we solve the negative-probability-density problem?

Recall from Eq.(23) that the Klein-Gordon equation produces negative probability density for negative energy. Apparently, the negative energy could not be avoided even for the Dirac equation. But *did we solve the negative-probability-density problem?* To answer this question, we need to obtain an expression for ρ and $\vec{j}$ which satisfies the continuity equation, just like we did for Eq.(16) and Eq.(22). Thus, let us start with the Dirac equation with the extended version of α_i and β (Eq.(43)):

$$
i\frac{\partial}{\partial t}\psi = (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)\psi = (\vec{\alpha}\cdot\vec{p} + \beta m)\psi \tag{46}
$$

where, now, ψ is a four-component vector, α_i and β are 4×4 matrices. From Eq.(32), we see that

$$
\sigma_i^\dagger = \sigma_i \tag{47}
$$

Thus, from Eq.(43), we also see that

$$
\alpha_i^\dagger = \alpha_i \qquad\qquad \beta^\dagger = \beta \tag{48}
$$

Take the Hermitian adjoint on both sides of Eq.(46):

$$
-i\frac{\partial}{\partial t}\psi^\dagger = \psi^\dagger(\overleftarrow{p}\cdot\vec{\alpha} + \beta m) \tag{49}
$$

where we used Eq.(48) to take Hermitian adjoint of α_i and β. Also, $\overleftarrow{p}$ means the momentum operator acting to the left. Then the Dirac equation and its Hermitian conjugate can be written as:

$$
\begin{aligned}
&\text{(i)} \qquad i\frac{\partial}{\partial t}\psi = (-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\psi \\
&\text{(ii)} \qquad -i\frac{\partial}{\partial t}\psi^\dagger = \psi^\dagger(i\vec{\alpha}\cdot\overleftarrow{\nabla} + \beta m)
\end{aligned}
\tag{50}
$$

Multiply $-i\psi^\dagger$ from the left of Eq.(50(i)) and $i\psi$ from the right of Eq.(50(ii))

$$
\begin{aligned}
&\text{(i)} \qquad \psi^\dagger\frac{\partial\psi}{\partial t} = -i\psi^\dagger(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\psi \\
&\text{(ii)} \qquad \frac{\partial\psi^\dagger}{\partial t}\psi = \psi^\dagger(i\vec{\alpha}\cdot\overleftarrow{\nabla} + \beta m)i\psi
\end{aligned}
\tag{51}
$$

By adding Eq.(51(i)) and Eq.(51.(ii)),

$$\psi^\dagger \frac{\partial \psi}{\partial t} + \frac{\partial \psi^\dagger}{\partial t} \psi = -i\psi^\dagger(-i\vec{\alpha} \cdot \vec{\nabla})\psi + \psi^\dagger(i\vec{\alpha} \cdot \overleftarrow{\nabla})i\psi \tag{52}$$

Notice that $\psi^\dagger \overleftarrow{\nabla} = \vec{\nabla}\psi^\dagger$. Thus, Eq.(52) becomes:

$$\frac{\partial}{\partial t}\left(\psi^\dagger \psi\right) = -\psi^\dagger \vec{\alpha} \cdot \left(\vec{\nabla}\psi\right) - \left(\vec{\nabla}\psi^\dagger\right) \cdot \vec{\alpha}\psi \tag{53}$$

We can combine the RHS of Eq.(53) as:

$$\frac{\partial}{\partial t}\left(\psi^\dagger \psi\right) = -\vec{\nabla} \cdot \left(\psi^\dagger \vec{\alpha} \psi\right) \tag{54}$$

Thus, the probability density and the current are defined as:

$$\text{(i)} \quad \rho = \psi^\dagger \psi \qquad\qquad \text{(ii)} \quad \vec{j} = \psi^\dagger \vec{\alpha} \psi \tag{55}$$

Eq.(55(i)) is always non-negative. So yes, we solved the negative-probability-density problem. For those who feel uncomfortable with $\overleftarrow{\nabla}$, let me give a quick justification. From linear algebra, we know

$$(A\vec{v})^\dagger = \vec{v}^\dagger A^\dagger$$

Let $A = -i\frac{d}{dx}$ and $\vec{v} = \psi$. Then

$$\left(-i\frac{d}{dx}\psi\right)^\dagger = \psi^\dagger \left(-i\frac{d}{dx}\right)^\dagger = \psi^\dagger \left(i\frac{d}{dx}\right) \tag{56}$$

Independently, we know that:

$$\left(-i\frac{d}{dx}\psi\right)^\dagger = i\frac{d}{dx}\psi^\dagger \tag{57}$$

By equating Eq.(56) and Eq.(57),

$$\psi^\dagger \left(i\frac{d}{dx}\right) = i\frac{d}{dx}\psi^\dagger \tag{58}$$

Thus, $\left(i\frac{d}{dx}\right)$ must act on $\psi^\dagger$ from the left:

$$\psi^\dagger \overleftarrow{\left(i\frac{d}{dx}\right)} = i\frac{d}{dx}\psi^\dagger \tag{59}$$

1.5.1 What *is* $\vec{\alpha}$?

In Eq.(55(ii)), the current of probability density is defined to be $\vec{j} = \psi^\dagger \vec{\alpha} \psi$. Then what does $\vec{\alpha}$ mean? To answer this question, let us recall Heisenberg picture. In quantum mechanics, an expectation value of an observable A at any time t is given by:

$$\langle A(t)\rangle = \langle \phi(t)| A |\phi(t)\rangle \tag{60}$$

where $\phi(t)$ is a state of an associated particle at any time t. Since $|\phi(t)\rangle = e^{-iHt}|\phi(0)\rangle$, Eq.(60) becomes:

$$\langle A(t)\rangle = \langle \phi(t)| A |\phi(t)\rangle = \langle \phi(0)| e^{iHt} A e^{-iHt} |\phi(0)\rangle \tag{61}$$

So we can say that an observable operator, A, does not change in time; however, a state ket, $|\phi(t)\rangle$ of a particle changes in time. On the other hand, we can equally claim that an observable operator $A(t) = e^{iHt}Ae^{-iHt}$ changes in time; however, a state ket, $|\phi(0)\rangle$, does not change in time. The former is called the Schrodinger picture, and the latter is called the Heisenberg picture.

Let us denote an operator of an observable A in Heisenberg picture to be A_H. Then its time evolution is given by:

$$\begin{aligned} \frac{d}{dt}A_H = \frac{d}{dt}\left(e^{iHt}Ae^{-iHt}\right) =& iHe^{iHt}Ae^{-iHt} + e^{iHt}A(-iH)e^{-iHt} \\ =& iHe^{iHt}Ae^{-iHt} + e^{iHt}Ae^{-iHt}(-iH) \\ =& iHA(t) - iA(t)H \\ =& i[H,\, A(t)] \end{aligned} \tag{62}$$

Let $A_H = \vec{x}(t)$ and $H = -\vec{\alpha}\cdot i\nabla + \beta m$. Then, by Eq.(62),

$$\frac{d\vec{x}}{dt} = i[-\vec{\alpha}\cdot i\nabla + \beta m,\, \vec{x}(t)] = i[-\vec{\alpha}\cdot i\nabla,\, \vec{x}(t)] = [\vec{\alpha}\cdot\nabla,\, \vec{x}(t)] \tag{63}$$

Let us act the commutator (63) on a position eigenket, $|\vec{x}'\rangle$, which satisfies $\vec{x}(t)|\vec{x}'\rangle = \vec{x}'(t)|\vec{x}'\rangle$, where $\vec{x}(t)$ is an operator and $\vec{x}'(t)$ is an eigenvalue of $|\vec{x}'\rangle$.

$$\begin{aligned} [\vec{\alpha}\cdot\nabla,\, \vec{x}(t)]|\vec{x}'\rangle =& \left(\vec{\alpha}\cdot\nabla\vec{x}(t) - \vec{x}(t)\vec{\alpha}\cdot\nabla\right)|\vec{x}'\rangle \\ =& \left(\vec{\alpha}\cdot\nabla\right)\left(\vec{x}(t)|\vec{x}'\rangle\right) - \vec{x}(t)\vec{\alpha}\cdot\nabla|\vec{x}'\rangle \\ =& \vec{\alpha}\cdot\left(\nabla'\vec{x}'(t)\right)|\vec{x}'\rangle + \vec{x}'(t)\vec{\alpha}\cdot\nabla'|\vec{x}'\rangle - \vec{x}'(t)\vec{\alpha}\cdot\nabla'|\vec{x}'\rangle \\ =& \vec{\alpha}\cdot\left(\nabla'\vec{x}'(t)\right)|\vec{x}'\rangle \\ =& \vec{\alpha}|\vec{x}'\rangle \end{aligned} \tag{64}$$

To explicitly lay out the calculations, let us focus on the second equality of Eq.(64).

$$\begin{aligned} &\left(\vec{\alpha}\cdot\nabla\right)\left(\vec{x}(t)|\vec{x}'\rangle\right) - \vec{x}(t)\vec{\alpha}\cdot\nabla|\vec{x}'\rangle \\ =&\left(\alpha_1\frac{\partial}{\partial x^1} + \alpha_2\frac{\partial}{\partial x^2} + \alpha_3\frac{\partial}{\partial x^3}\right)\left(x^1, x^2, x^3\right)|\vec{x}'\rangle \\ &-\left(x^1, x^2, x^3\right)\left(\alpha_1\frac{\partial}{\partial x^1} + \alpha_2\frac{\partial}{\partial x^2} + \alpha_3\frac{\partial}{\partial x^3}\right)|\vec{x}'\rangle \end{aligned} \tag{65}$$

Let's calculate the first component of the first term in Eq.(65):

$$\begin{aligned}
&\left(\alpha_1 \frac{\partial}{\partial x'^1} + \alpha_2 \frac{\partial}{\partial x'^2} + \alpha_3 \frac{\partial}{\partial x'^3}\right)\left(x'^1 \left|\vec{x}'\right\rangle\right) \\
=&\alpha_1 \left|\vec{x}'\right\rangle + x'^1 \left(\alpha_1 \frac{\partial}{\partial x'^1} + \alpha_2 \frac{\partial}{\partial x'^2} + \alpha_3 \frac{\partial}{\partial x'^3}\right)\left|\vec{x}'\right\rangle \\
=&\alpha_1 \left|\vec{x}'\right\rangle + x'^1 \left(\vec{\alpha} \cdot \nabla\right)\left|\vec{x}'\right\rangle
\end{aligned} \tag{66}$$

By considering all three components, the last line of Eq.(66) becomes:

$$\vec{\alpha} \left|\vec{x}'\right\rangle + \vec{x}' \left(\vec{\alpha} \cdot \nabla\right)\left|\vec{x}'\right\rangle \tag{67}$$

Then the second term in Eq.(67) is canceled out by the second term in the second equality of Eq.(64), which results in $\vec{\alpha}\left|\vec{x}'\right\rangle$. Thus, from Eq.(64),

$$\frac{d\vec{x}}{dt} = [\vec{\alpha} \cdot \nabla, \vec{x}(t)] = \vec{\alpha} \tag{68}$$

Hence, $\vec{\alpha}$ *is* a velocity operator. In a classical sense, this sounds reasonable. Restating Eq.(55),

$$\text{(i)} \quad \rho = \psi^\dagger \psi \qquad\qquad \text{(ii)} \quad \vec{j} = \psi^\dagger \vec{\alpha} \psi$$

Since $\frac{d\vec{x}}{dt} = \vec{\alpha}$, we can say $\vec{j}$ corresponds to $\rho\vec{v}$.

1.6 Spin of a Particle

Before jumping into how spin is derived from the Dirac equation, we shall discuss some mathematical properties.

1.6.1 $[A,\, BC]$, $[AB,\, C]$, and $[AB,\, CD]$

Let A, B, and C be operators. We shall expand $[A,\, BC]$, $[AB,\, C]$, and $[AB,\, CD]$.

$$\begin{aligned}
[A,\, BC] =& ABC - BCA = ABC - BCA + BAC - BAC \\
=& ABC - BAC + BAC - BCA = (AB - BA)C + B(AC - CA) \\
=& [A,\, B]C + B[A,\, C]
\end{aligned}$$

$$\therefore [A,\, BC] = [A,\, B]C + B[A,\, C] \tag{69}$$

Next,

$$\begin{aligned}
[AB,\, C] =& ABC - CAB = ABC - CAB + ACB - ACB \\
=& ABC - ACB + ACB - CAB = A(BC - CB) + (AC - CA)B \\
=& A[B,\, C] + [A,\, C]B
\end{aligned}$$

$$\therefore [AB,\, C] = A[B,\, C] + [A,\, C]B \tag{70}$$

Finally, we need to use Eq.(69) and Eq.(70) to expand the last commutator.

$$\begin{aligned}
[AB,\, CD] =& [AB,\, C]D + C[AB,\, D] \\
=& A[B,\, C]D + [A,\, C]BD + CA[B,\, D] + C[A,\, D]B
\end{aligned}$$

$$\therefore [AB,\, CD] = A[B,\, C]D + [A,\, C]BD + CA[B,\, D] + C[A,\, D]B \tag{71}$$

1.6.2 Measuring Energy and Spin Simultaneously

Is it possible to measure energy (spin) state without destroying spin (energy) state? To answer this question, we need to calculate the commutator of the Hamiltonian of a particle and its angular momentum operator.

$$[H,\, \vec{J}] \tag{72}$$

where

$$H = \vec{\alpha}\cdot\vec{p} + \beta m \qquad \vec{J} = \vec{L} + \vec{S} \qquad \vec{L} = \vec{x}\times\vec{p} \qquad S_i = \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} \tag{73}$$

Here, S_i is a i-th component of $\vec{S}$. It is a 4×4 generalization of spin matrices (just like we extended α_i into 4×4 dimension). Using Eq.(69), (70), and (71),

$$\begin{aligned}
[H,\, J_i] =&[\vec{\alpha}\cdot\vec{p} + \beta m,\, L_i + S_i] = [\alpha_l p_l + \beta m,\, \epsilon_{ijk}x_j p_k + S_i]\\
=&[\alpha_l p_l,\, \epsilon_{ijk}x_j p_k] + [\alpha_l p_l,\, S_i] + [\beta m,\, \epsilon_{ijk}x_j p_k] + [\beta m,\, S_i]\\
=&\epsilon_{ijk}[\alpha_l p_l,\, x_j p_k] + [\alpha_l p_l,\, S_i] + m\epsilon_{ijk}[\beta,\, x_j p_k] + m[\beta,\, S_i]\\
=&\epsilon_{ijk}(\alpha_l[p_l,\, x_j]p_k + [\alpha_l,\, x_j]p_l p_k + x_j\alpha_l[p_l,\, p_k] + x_j[\alpha_l,\, p_k]p_l)\\
&+ \alpha_l[p_l,\, S_i] + [\alpha_l,\, S_i]p_l + m\epsilon_{ijk}([\beta,\, x_j]p_k + x_j[\beta,\, p_k]) + m[\beta,\, S_i]
\end{aligned} \tag{74}$$

Let us inspect each commutator:

$$[p_l,\, x_j] = -i\delta_{lj} \qquad [\alpha_l \text{ or } \beta,\, x_j \text{ or } p_k] = 0 \qquad [p_l,\, p_k] = 0 \qquad [p_l,\, S_i] = 0 \tag{75}$$

Then the last line of Eq.(74) becomes:

$$[H,\, J_i] = -i\epsilon_{ijk}\alpha_l\delta_{lj}p_k + [\alpha_l,\, S_i]p_l + m[\beta,\, S_i] \tag{76}$$

$$\begin{aligned}
[\alpha_l,\, S_i] = \alpha_l S_i - S_i\alpha_l =&\frac{1}{2}\begin{pmatrix}0 & -\sigma_l\\ \sigma_l & 0\end{pmatrix}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} - \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}\begin{pmatrix}0 & -\sigma_l\\ \sigma_l & 0\end{pmatrix}\\
=&\frac{1}{2}\begin{pmatrix}0 & -\sigma_l\sigma_i\\ \sigma_l\sigma_i & 0\end{pmatrix} - \frac{1}{2}\begin{pmatrix}0 & -\sigma_i\sigma_l\\ \sigma_i\sigma_l & 0\end{pmatrix}\\
=&\frac{1}{2}\begin{pmatrix}0 & -[\sigma_l,\, \sigma_i]\\ [\sigma_l,\, \sigma_i] & 0\end{pmatrix}\\
=&\frac{1}{2}\begin{pmatrix}0 & -2i\epsilon_{lik}\sigma_k\\ 2i\epsilon_{lik}\sigma_k & 0\end{pmatrix} = i\epsilon_{lik}\begin{pmatrix}0 & -\sigma_k\\ \sigma_k & 0\end{pmatrix}\\
=&i\epsilon_{lik}\alpha_k
\end{aligned} \tag{77}$$

$$\begin{aligned}
[\beta,\, S_i] =&\beta S_i - S_i\beta\\
=&\frac{1}{2}\begin{pmatrix}0 & I_2\\ I_2 & 0\end{pmatrix}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} - \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}\begin{pmatrix}0 & I_2\\ I_2 & 0\end{pmatrix}\\
=&\frac{1}{2}\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix} - \frac{1}{2}\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix} = 0
\end{aligned} \tag{78}$$

Thus, by plugging the results of Eq.(77) and Eq.(78) into Eq.(76):

$$\begin{aligned}[H,\, J_i] &= -\, i\epsilon_{ijk}\alpha_l\delta_{lj}p_k + i\epsilon_{lik}\alpha_k p_l \\ &= -\, i\epsilon_{ijk}\alpha_j p_k + i\epsilon_{lik}\alpha_k p_l \\ &= -\, i\epsilon_{ijk}\alpha_j p_k + i\epsilon_{ikl}\alpha_k p_l = 0\end{aligned} \tag{79}$$

Hence, the Hamiltonian and its total angular momentum operator commute. This implies that there exists a set of states which are common eigenstates of H and $\vec{J}$. Let us take a special case where $\vec{p} = 0$, which means that we are in the frame stationary relative to a particle under consideration. Then

$$H = \beta m \qquad\qquad J_i = S_i \tag{80}$$

This means there exist some states which a particle's rest energy and its spin can be measured at the same time.

1.6.3 Spin-1/2

In Section 1.4, we let $m = 0$. This time, we let $\vec{p} = 0$, which means that we are in the frame stationary relative to the particle. If we choose an energy eigenstate ψ, then the Dirac equation becomes:

$$E\psi = \beta m\psi = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} m\psi = \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \psi \tag{81}$$

Keep in mind that m is a 2×2 matrix, which is $I_2 m$. Since two 2-component vectors, Eq.(81) can be written as:

$$E \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rightarrow \begin{pmatrix} E\phi_1 \\ E\phi_2 \end{pmatrix} = \begin{pmatrix} m\phi_2 \\ m\phi_1 \end{pmatrix} \tag{82}$$

Since $\vec{p} = 0$, $E^2 = |\vec{p}|^2 + m^2 = m^2$, which implies $E = \pm m$.

(i) Suppose $E = +m$.

Then, from Eq.(82), we have:

$$\begin{pmatrix} E\phi_1 \\ E\phi_2 \end{pmatrix} = \begin{pmatrix} m\phi_2 \\ m\phi_1 \end{pmatrix} \rightarrow \phi_1 = \phi_2 \tag{83}$$

Since ϕ_1 and ϕ_2 are 2-component vectors, we can let them to be either

$$\phi_1 = \phi_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{or} \qquad \phi_1 = \phi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{84}$$

Thus, the 4-component state is:

$$\psi_{+E,\uparrow} = \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \quad \text{or} \quad \psi_{+E,\downarrow} = \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \tag{85}$$

(ii) Suppose $E = -m$.

Then, from Eq.(82), we have:

$$\begin{pmatrix} E\phi_1 \\ E\phi_2 \end{pmatrix} = \begin{pmatrix} m\phi_2 \\ m\phi_1 \end{pmatrix} \rightarrow \phi_1 = -\phi_2 \tag{86}$$

Since ϕ_1 and ϕ_2 are 2-component vectors, we can let them to be either

$$\phi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \phi_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{or} \quad \phi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \phi_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{87}$$

Thus, the 4-component state is:

$$\psi_{-E,\uparrow} = \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{pmatrix} \quad \text{or} \quad \psi_{-E,\downarrow} = \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \\ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{pmatrix} \tag{88}$$

Now we apply S_3 on $\psi_{+E,\uparrow}$, $\psi_{+E,\downarrow}$, $\psi_{-E,\uparrow}$, and $\psi_{-E,\downarrow}$:

$$S_3\psi_{+E,\uparrow} = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \frac{1}{2}\psi_{+E,\uparrow} \tag{89}$$

$$S_3\psi_{+E,\downarrow} = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = -\frac{1}{2}\psi_{-E,\downarrow} \tag{90}$$

$$S_3\psi_{-E,\uparrow} = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{pmatrix} = \frac{1}{2}\psi_{-E,\uparrow} \tag{91}$$

$$S_3\psi_{-E,\downarrow} = \frac{1}{2}\begin{pmatrix}\sigma_3 & 0 \\ 0 & \sigma_3\end{pmatrix}\begin{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \\ \\ \begin{pmatrix}0\\-1\end{pmatrix}\end{pmatrix} = -\frac{1}{2}\begin{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \\ \\ \begin{pmatrix}0\\-1\end{pmatrix}\end{pmatrix} = -\frac{1}{2}\psi_{-E,\downarrow} \tag{92}$$

Eq.(89) and Eq.(91) represent spin-1/2 up, and Eq.(90) and Eq.(92) represent spin-1/2 down state. Thus, the Dirac equation describes a spin-1/2 particle! The Dirac equation derives the spin of a particle, which the Schrodinger equation could not. However, there is still a freedom to choose the sign of the energy.

1.7 Interpretation of Negative Energy

What does negative energy mean? Here, we present Dirac's interpretation first and then Feynman-Stuckelberg interpretation later.

1.7.1 Dirac's Hole Theory

Suppose a particle has negative energy. Then it can emit energy over and over to reduce its energy level, resulting in releasing an infinite amount of energy. Indeed, it does not physically make sense. Thus, Dirac thought that all of the negative energy levels are already filled by electrons. Then an electron with negative energy cannot occupy a lower energy level (which is already filled by two electrons with opposite spins) due to the Pauli exclusion principle. The idea that electrons have already occupied all the negative energy levels is called the Dirac sea, and Dirac believed that this sea is vacuum!

Let's say there is an electron with $E = -mc^2$. If this electron obtains at least $E = +2mc^2$, it has an energy of $E = +mc^2$. Then a "hole" is created in the energy level of $E = -mc^2$. This hole moves in the opposite direction to where electrons in the Dirac sea move when an external electric field is applied. Thus, the hole is like an electron with a positive charge. Dirac predicted the existence of antiparticles. One problem is, this reasoning does not work for bosons, for which the Pauli exclusion principle is not applied.

1.7.2 Feynman-Stuckelberg Interpretation

Let us revive the Klein-Gordon equation and focus on Eq.(22). By taking

$$\frac{\partial}{\partial t} \to \partial_0 = \partial^0 \qquad\qquad \frac{\partial}{\partial x^i} \to \partial_i = -\partial^i$$

Eq.(22) can be combined and rewritten as:

$$j^\mu = i(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*) \tag{93}$$

where

$$\rho = j^0 = i(\varphi^* \partial^0 \varphi - \varphi \partial^0 \varphi^*) \qquad\qquad \vec{j} = j^i = i(\varphi^* \partial^i \varphi - \varphi \partial^i \varphi^*)$$

With a plane-wave solution $\varphi = Ae^{-i(Et - \vec{p}\cdot\vec{x})}$ $(E > 0)$, j^μ is defined by:

$$j^\mu = 2|A|^2 (E, \vec{p}) \tag{94}$$

Let us view j^0 as a charge-current density instead of a probability-current density. So we multiply $-e$ $(e > 0)$ on the RHS side of Eq.(94).

$$j^\mu(-e) = -2e|A|^2 (E, \vec{p}) \tag{95}$$

By adopting the view that j^μ is a charge-current density, we do not have to bother with the fact that j^0 can be negative. Let us take the time-reversal operation:

$$t \to -t \qquad\qquad \vec{p} = m\frac{d\vec{x}}{dt} \to m\frac{d\vec{x}}{d(-t)} = -\vec{p} \tag{96}$$

Then the time-reversed plane-wave solution becomes: $\varphi = Ae^{i(Et - \vec{p}\cdot\vec{x})}$. The net effect is $\vec{p} \to -\vec{p}$ and $E \to -E$. Then its probability-current density (Eq.(94)) is modified into:

$$j^\mu = 2|A|^2 (-E,\ -\vec{p}) \qquad (E > 0) \tag{97}$$

Then again, by multiplying $-e$ on the RHS of Eq.(97), the charge-current density,

$$j^\mu(-e) = -2e|A|^2 (-E,\ -\vec{p}) \qquad (E > 0) \tag{98}$$

Eq.(98) is a time-reversed version of Eq.(95), because it comes from the time-reversed plane-wave solution. However, Eq.(98) can also be obtained by simply taking $-e \to +e$ in Eq.(95):

$$j^\mu(+e) = +2e|A|^2 (E, \vec{p}) = -2e|A|^2 (-E,\ -\vec{p}) \qquad (E > 0) \tag{99}$$

Thus, combining Eq.(98) and Eq.(99),

$$j^\mu(+e)_{time-forwarded} = j^\mu(-e)_{time-reversed} \tag{100}$$

Therefore, we claim the following: *The negative energy of a particle corresponds to the particle traveling backward in time, and this is the same as a particle with an opposite charge (a.k.a. antiparticle) traveling forward in time.*

1.8 Does the Dirac equation keep its form under the Lorentz transformation?

Laws of physics have to be invariant no matter how fast the frame is moving relative to another. The section discusses if that is true for the Dirac equation.

1.9 γ Matrices and their Properties

Let us start with the Dirac equation:

$$i\frac{\partial}{\partial t}\psi = (\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)\psi$$

Multiply β from the left.

$$i\beta\frac{\partial}{\partial t}\psi = \beta(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)\psi$$

Let us define following:

$$\beta = \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad \beta\alpha_i = \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{101}$$

You might feel uncomfortable about $\beta\alpha_i = \gamma^i$ because the index is subscripted on the LHS and superscripted on the RHS. However, (β, α_1, α_2, α_3) is **not** a 4-vector, which means that they do not obey the Lorentz transformation. Hence, the indices on γ^i and γ_i are labels for the matrices and do not imply 4-vector transformation properties.

With the substitution (101), the Dirac equation is written as:

$$i\gamma^0\frac{\partial}{\partial t}\psi = (\gamma^1 p_1 + \gamma^2 p_2 + \gamma^3 p_3 + \beta^2 m)\psi \tag{102}$$

Since

$$p_i = -i\frac{\partial}{\partial x^i} = -i\partial_i \qquad\qquad \beta^2 = I$$

Eq.(102) is:

$$i\gamma^0\frac{\partial}{\partial t}\psi = (-i\gamma^1\partial_1 - i\gamma^2\partial_2 - i\gamma^3\partial_3 + m)\psi$$

$$i\gamma^0\partial_0\psi + (i\gamma^1\partial_1 + i\gamma^2\partial_2 + i\gamma^3\partial_3 - m)\psi = 0$$

$$(\gamma^\mu i\partial_\mu - m)\psi = 0 \tag{103}$$

We had the Dirac conditions (Eq.(29)) for α_i and β. Let us see what conditions the Dirac matrices γ^μ must satisfy. Originally,

$$\text{(i)} \quad \alpha_i^2 = \beta^2 = I, \quad \text{(ii)} \quad \{\alpha_i, \alpha_j\} = 0 \quad \text{if } i \neq j, \quad \text{(iii)} \quad \{\alpha_i, \beta\} = 0$$

From condition (i) and Eq.(101),

$$\beta^2 = I \rightarrow (\gamma^0)^2 = I \tag{104}$$

and with condition (iii)

$$(\gamma^i)^2 = (\beta\alpha_i)(\beta\alpha_i) = \beta(\alpha_i\beta)\alpha_i = -\beta(\beta\alpha_i)\alpha_i = -\beta^2\alpha_i^2 = -I$$
$$\rightarrow (\gamma^i)^2 = -I \tag{105}$$

From condition (ii),

$$\{\alpha_i, \alpha_j\} = 0 \rightarrow \alpha_i\alpha_j + \alpha_j\alpha_i = 0$$

By multiplying β from the left and the right,

$$\beta\alpha_i\alpha_j\beta + \beta\alpha_j\alpha_i\beta = 0$$

Due to condition (iii), $\alpha_i\beta = -\beta\alpha_i$. The equation above can be written as:

$$\beta\alpha_i(-\beta\alpha_j) + \beta\alpha_j(-\beta\alpha_i) = 0$$

Thus, by recalling Eq.(101),

$$\gamma_i\gamma_j + \gamma_j\gamma_i = \{\gamma_i,\, \gamma_j\} = 0 \qquad (i \neq j) \tag{106}$$

From condition (iii),

$$\{\alpha_i, \beta\} = 0 \rightarrow \alpha_i\beta + \beta\alpha_i = 0$$

Multiply β from the left.

$$\beta\alpha_i\beta + \beta\beta\alpha_i = 0$$

By recalling Eq.(101) again,

$$\gamma^0\gamma^i + \gamma^i\gamma^0 = \{\gamma^i,\, \gamma^0\} = 0 \tag{107}$$

Restate all the conditions for the Dirac matrices (Eq.(104), Eq.(105), Eq.(106), and Eq.(107)):

$$\begin{array}{ll} \text{(i)} \quad (\gamma^0)^2 = I, \quad \text{(ii)} \quad (\gamma^i)^2 = -I \\ \text{(iii)} \quad \{\gamma^i,\, \gamma^j\} = 0 \quad \text{if } i \neq j, \quad \text{(iv)} \quad \{\gamma^i,\, \gamma^0\} = 0 \end{array} \tag{108}$$

We can combine all the conditions in Eq.(108):

$$\{\gamma^\mu,\, \gamma^\nu\} = 2g^{\mu\nu} \tag{109}$$

Also, since

$$\beta = \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \beta\alpha_i = \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \qquad \sigma_i^\dagger = \sigma_i \tag{110}$$

The first two equations are just from Eq.(101). σ_i must be Hermitian because σ_i is an observable as a spin value, omitting $\hbar/2$. Then, from Eq.(110), you can see that

$$\gamma^{0\dagger} = \gamma^0 \qquad\qquad \gamma^{i\dagger} = -\gamma^i \tag{111}$$

Therefore, taking the Hermitian conjugate of the gamma matrices is like raising/lowering indices on $(\gamma^0, \vec{\gamma})$! Eq.(109) and Eq.(111) are called the Dirac conditions for gamma matrices.

$$\text{(i)} \quad \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \qquad \text{(ii)} \quad \gamma^{0\dagger} = \gamma^0, \qquad \text{(iii)} \quad \gamma^{i\dagger} = -\gamma^i \tag{112}$$

Before proceeding any further, we wish to prove one more identity (it is not too bad, please bear with me), which is:

$$\gamma^\mu\gamma^0 - \gamma^0\gamma^{\mu\dagger} = 0 \tag{113}$$

if $\mu = 0$, then $\quad \gamma^\mu\gamma^0 - \gamma^0\gamma^{\mu\dagger} = \gamma^0\gamma^0 - \gamma^0\gamma^{0\dagger} = \gamma^0\gamma^0 - \gamma^0\gamma^0 = 0$

if $\mu = k$, then $\quad \gamma^\mu\gamma^0 - \gamma^0\gamma^{\mu\dagger} = \gamma^k\gamma^0 - \gamma^0\gamma^{k\dagger} = \gamma^k\gamma^0 - \gamma^0(-\gamma^k) = \{\gamma^k, \gamma^0\} = 0$

where we used Eq.(112). Thus, identity (113) is proved. Then

$$\gamma^\mu\gamma^0 - \gamma^0\gamma^{\mu\dagger} = 0 \to \gamma^0\gamma^{\mu\dagger} = \gamma^\mu\gamma^0 \to \gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu \tag{114}$$

where we used $(\gamma^0)^2 = I$. Let's just include Eq.(114) into the Dirac conditions Eq.(112):

$$\begin{aligned} &\text{(i)} \quad \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \qquad \text{(ii)} \quad \gamma^{0\dagger} = \gamma^0, \\ &\text{(iii)} \quad \gamma^{i\dagger} = -\gamma^i, \qquad \text{(iv)} \quad \gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu \end{aligned} \tag{115}$$

1.10 $\bar{\psi}\gamma^\mu\psi$ seems like a 4-vector!

Let us go back to the Dirac equation (Eq.(103)):

$$(\gamma^\mu i\partial_\mu - m)\psi = 0$$

Take the Hermitian conjugate on both sides:

$$\psi^\dagger([\gamma^\mu i\partial_\mu]^\dagger - m) = \psi^\dagger([\gamma^0 i\partial_0]^\dagger + [\gamma^k i\partial_k]^\dagger - m) = 0 \tag{116}$$

Recall Eq.(115(ii)), Eq.(115(iii)), $(i\partial_0)^\dagger = -i\partial_0$, and $(i\partial_k)^\dagger = i\partial_k$. ($i\partial_k$ must be Hermitian because it is just a minus of the momentum operator, which is Hermitian.) But this is **not** case this time! Eq.(115(ii)) and Eq.(115(iii)) hold, but there is a problem with $(i\partial_0)^\dagger = -i\partial_0$ and $(i\partial_k)^\dagger = i\partial_k$. Let's briefly talk about why.

To discuss hermiticity, we must first identify the vector space in question. Here, two spaces are involved: the 4-dimensional spinor space and the infinite-dimensional Hilbert space of square-integrable functions. The derivative operator ∂_k is anti-Hermitian (meaning $i\partial_k$ is Hermitian) in the Hilbert space, but not in the spinor space. The Hermitian conjugate of the Dirac equation is applied in the spinor space in our derivation. To avoid confusion: Lorentz spinors are not Lorentz vectors, so how can spinors belong to a vector space? Here's a clarification: 4-dimensional spacetime is distinct from the n-dimensional space where the SO(1,3) group acts. A spinor is not a vector in spacetime but is a vector in the 4-dimensional spinor space for the Dirac equation. Thus, ∂_μ remains the same even after taking its Hermitian conjugate in the spinor space. Hence, in spinor space,

$$\gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu, \qquad (\partial_\mu)^\dagger = \partial_\mu, \qquad i^\dagger = -i$$

Then Eq.(116) is written as:

$$\psi^\dagger([\gamma^0 i\partial_0]^\dagger + [\gamma^k i\partial_k]^\dagger - m) = \psi^\dagger(-(\gamma^0)^\dagger i\partial_0 - (\gamma^k)^\dagger i\partial_k - m) = 0 \tag{117}$$

Since $\gamma^0\gamma^0 = I$, we can plug it in any place we want. Also, we can multiply γ^0 on both sides of Eq.(117) from the right:

$$\begin{aligned}
&\psi^\dagger \gamma^0\gamma^0(-(\gamma^0)^\dagger i\partial_0 - (\gamma^k)^\dagger i\partial_k - m)\gamma^0 \\
=&(\psi^\dagger\gamma^0)(-\gamma^0(\gamma^0)^\dagger\gamma^0 i\partial_0 - \gamma^0(\gamma^k)^\dagger\gamma^0 i\partial_k - \gamma^0 m\gamma^0) \\
=&(\psi^\dagger\gamma^0)(-(\gamma^0(\gamma^\mu)^\dagger\gamma^0) i\partial_\mu - m\gamma^0\gamma^0) \\
=&-\bar{\psi}(\gamma^\mu i\partial_\mu + m) = 0
\end{aligned} \tag{118}$$

where $\bar{\psi} = \psi^\dagger\gamma^0$, called the adjoint of ψ. On the last line, the partial derivative ∂_μ acts from the right on $\bar{\psi}$. Hence,

$$\bar{\psi}(\gamma^\mu i\overleftarrow{\partial_\mu} + m) = 0 \tag{119}$$

We can also derive a new version of the continuity equation (Eq.(54)). Let us begin with Eq.(54):

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) = -\vec{\nabla}\cdot\left(\psi^\dagger\vec{\alpha}\psi\right) = -\partial_k\left(\psi^\dagger\alpha^k\psi\right)$$

Again, we can plug $\gamma^0\gamma^0 = I$ into anywhere we want.

$$\frac{\partial}{\partial t}\left(\psi^\dagger\gamma^0\gamma^0\psi\right) = -\partial_k\left(\psi^\dagger\alpha^k\psi\right) = -\partial_k\left(\psi^\dagger\gamma^0\gamma^0\alpha^k\psi\right) \tag{120}$$

Recall:

$$\gamma^0\alpha_i = \gamma^i \qquad \bar{\psi} = \psi^\dagger\gamma^0$$

So Eq.(120) is written as:

$$\frac{\partial}{\partial t}\left(\psi^\dagger\gamma^0\gamma^0\psi\right)+\partial_k\left(\psi^\dagger\gamma^0\gamma^0\alpha^k\psi\right)=\frac{\partial}{\partial t}\left(\bar{\psi}\gamma^0\psi\right)+\partial_k\left(\bar{\psi}\gamma^k\psi\right)=0$$

Hence the last line can be written as:

$$\partial_\mu(\bar{\psi}\gamma^\mu\psi)=0 \tag{121}$$

The form of Eq.(121) suggests that $\bar{\psi}\gamma^\mu\psi$ is a 4-vector. So maybe it obeys the Lorentz transformation law. Then, a natural question arises: what transformation law does $\psi(x)$ obey?

1.11 Transformation Law of ψ

What *is* a vector? Is it just a row or column array of numbers? Well that's what a vector *looks* like, but it is not a correct definition. Generally, a vector (or tensor if you like) is defined as a geometrical entity which obeys a certain transformation rule. So we see two somewhat abstract ideas: (1) a geometrical entity and (2) a certain transformation rule. A geometrical entity means a vector exists even without defining a coordinate system. Its components are coordinate-dependent, but its presence is not affected by which coordinate you are in. However, this is not what we focus on in this section. Rather we want to discuss a certain transformation rule that ψ has to follow.

1.11.1 Rotation of a 2D-Euclidean Vector

One of the most well-known examples of a vector transformation is the rotational transformation of a 2D-Euclidean vector. Suppose we have a vector $\vec{v}=x\hat{x}+y\hat{y}$ in the Cartesian coordinate. We want to rotate this vector around the origin (or about the z-axis) by θ in the counterclockwise direction. Then the result is quite well-known:

$$\vec{v}'=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\cos\theta-y\sin\theta\\ x\sin\theta+y\cos\theta\end{pmatrix} \tag{122}$$

Thus we can say $\vec{v}=x\hat{x}+y\hat{y}$ follows a certain transformation rule (which is rotation here) and therefore $\vec{v}$ is a vector.

1.11.2 Another way to write the rotation matrix

In Section 1.9.1, we had a rotation matrix which rotates a 2D Euclidean vector by an angle θ about the origin. There is another way to write the rotation matrix, and its form is going to appear frequently in future discussion.

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\cos\theta\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\sin\theta\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \tag{123}$$

Let

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{124}$$

Notice

$$A^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I \tag{125}$$

Therefore

$$A^3 = A^2 A = -IA = -A \qquad\qquad A^4 = A^2 A^2 = (-I)(-I) = I \tag{126}$$

Back to Eq.(123) with the substitution (124):

$$I\cos\theta + A\sin\theta = I\left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + A\left(\frac{\theta^1}{1!} - \frac{\theta^3}{3!} + \cdots\right) \tag{127}$$

We can insert $I = -A^2$ or $I = A^4$ wherever we want. Thus

$$\begin{aligned} I\cos\theta + A\sin\theta =& I\left(1 - (-A^2)\frac{\theta^2}{2!} + (A^4)\frac{\theta^4}{4!} - \cdots\right) + A\left(\frac{\theta^1}{1!} - (-A^2)\frac{\theta^3}{3!} + \cdots\right) \\ =& I\left(1 + \frac{(A\theta)^2}{2!} + \frac{(A\theta)^4}{4!} + \cdots\right) + \left(\frac{(A\theta)^1}{1!} + \frac{(A\theta)^3}{3!} + \cdots\right) \\ =& I + \frac{(A\theta)^1}{1!} + \frac{(A\theta)^2}{2!} + \frac{(A\theta)^3}{3!} + \frac{(A\theta)^4}{4!} + \cdots = e^{A\theta} \end{aligned} \tag{128}$$

We just expressed the rotation matrix $R_z(\theta)$ as $e^{A\theta}$. But we shouldn't be too surprised that a rotation can be expressed in an exponential form.

Let us consider $e^{i\theta}$. Any complex number can be expressed as $Ne^{i\phi}$, where N is a real number. Let us multiply these two together:

$$e^{i\theta} \cdot Ne^{i\phi} = Ne^{i(\theta+\phi)}$$

In the complex plane, this represents rotating a phase ϕ by θ. Thus, an exponential form can be used to express a rotation operator.

1.11.3 Transformation of a Vector Function

What *is* a vector function? It is a machine which takes a vector in one space and produces another vector in the same or different vector space. The key thing to note is that the input vector and the output vector do not have to be in the same vector space. Hence, the input and output vectors can have different transformation rules, because they may belong to different vector spaces.

Let us denote the transformation rule of input vectors and output vectors as:

$$\vec{x}' = R\vec{x} \qquad\qquad \vec{V}' = O(R)\vec{V} \tag{129}$$

where $O(R)$ stands for the transformation *operator* of output vectors, associated with the transformation rule R of input vectors. *All* that Eq.(129) says is that if an input vector changes in a certain way, then the output vector also changes in a corresponding way. Again, notice that R and $O(R)$ are two different operators since $\vec{x}$ and $\vec{V}$ can be in two different vector spaces! Then the following relation holds:

$$\vec{V}'(\vec{x}) = O(R)\vec{V}(R^{-1}\vec{x}) \tag{130}$$

Here is what Eq.(130) means. Suppose you have a board on which there is a mark on every point. We call an arbitrary mark as $\vec{x}$. Also, assume that there is a super tiny arrow assigned on each mark. We call this arrow on a mark $\vec{x}$ to be $\vec{V}(\vec{x})$. Suppose you are looking at two marks $\vec{x}$ and $R^{-1}\vec{x}$, where $\vec{x}$ can be obtained by rotating $R^{-1}\vec{x}$ by a certain angle. Then you see $\vec{V}(\vec{x})$ and $\vec{V}(R^{-1}\vec{x})$ are assigned on $\vec{x}$ and $R^{-1}\vec{x}$, respectively. Now you rotate the whole set of arrows $\vec{V}$ by applying $O(R)$ on each arrow *without* rotating the board. We call the set of rotated arrows to be $\vec{V}'$. But you are still looking at the two marks $\vec{x}$ and $R^{-1}\vec{x}$. Then now you see that the new vector on $\vec{x}$, which is $\vec{V}'(\vec{x})$ came from $\vec{V}(R^{-1}\vec{x})$, but rotated by $O(R)$!

By setting $\vec{x} \to R\vec{x}$, we can rewrite Eq.(130) to be:

$$\vec{V}'(\vec{x}') = O(R)\vec{V}(\vec{x}) \tag{131}$$

where $\vec{x}' = R\vec{x}$.

In our context, the mark is a point in spacetime and the arrow is a spinor in its spinor space. In other words, the machine takes a vector in space-time whose coordinate is x and produces $\psi(x)$ which lives in a spinor space:

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \to \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \tag{132}$$

where each ψ_i is a complex number.

Thus we want to find the transformation operators L and $O(L)$ such that

$$x' = Lx \qquad\qquad \psi'(x') = O(L)\psi \tag{133}$$

Thankfully, we already know how x is transformed.

1.11.4 Spatial Rotation of x

Spatial rotation of x affects only the spatial components of x, which are x^1, x^2, and x^3. Let us denote each rotation matrix about the i-th axis to be R_i. Then:

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \quad R_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

If the angle θ is infinitesimally small, then $\theta \to \delta\theta$. R_i becomes:

$$R_1(\delta\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\delta\theta \\ 0 & \delta\theta & 1 \end{pmatrix} \quad R_2(\delta\theta) = \begin{pmatrix} 1 & 0 & \delta\theta \\ 0 & 1 & 0 \\ -\delta\theta & 0 & 1 \end{pmatrix} \quad R_3(\delta\theta) = \begin{pmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Denote $\vec{x} = (x^1, x^2, x^3)$ and let us apply $R_i(\delta\theta)$ on $\vec{x}$.

$$R_1(\delta\theta)\vec{x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\delta\theta \\ 0 & \delta\theta & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 - x^3\delta\theta \\ x^2\delta\theta + x^3 \end{pmatrix}$$

$$R_2(\delta\theta)\vec{x} = \begin{pmatrix} 1 & 0 & \delta\theta \\ 0 & 1 & 0 \\ -\delta\theta & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 + x^3\delta\theta \\ x^2 \\ -x^1\delta\theta + x^3 \end{pmatrix} \tag{134}$$

$$R_3(\delta\theta)\vec{x} = \begin{pmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 - x^2\delta\theta \\ x^1\delta\theta + x^2 \\ x^3 \end{pmatrix}$$

1.11.5 Lorentz Transformation of x

Other than spatial rotation, x can be Lorentz transformed. Suppose a primed coordinate system is moving at speed v into the x-direction relative to an unprimed coordinate system.

$$\begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \gamma(x^0 + \beta x^1) \\ \gamma(\beta x^0 + x^1) \\ x^2 \\ x^3 \end{pmatrix} \tag{135}$$

Also, for a given point in space-time, we can define space-time invariant. For one-dimensional motion, the space-time invariant is given by:

$$s^2 = (x^0)^2 - (x^1)^2 \tag{136}$$

Because space-time invariant is constant in all frames, we can define x^0 and x^1 in the following:

$$x^0 = R\cosh\eta \qquad\qquad x^1 = R\sinh\eta \tag{137}$$

Suppose x^0 and x^1 track when/where the moving frame is. Assuming that the two frames coincided at $x^0 = 0$, the velocity of the moving frame is given by:

$$v = \beta = \frac{x^1}{x^0} = \frac{R\sinh\eta}{R\cosh\eta} = \tanh\eta$$

Thus,

$$\beta = \tanh \eta \tag{138}$$

Then γ is given by:

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\tanh^2 \eta}} = \cosh \eta \tag{139}$$

Then $\beta\gamma$ is given by:

$$\beta\gamma = \tanh \eta \cdot \cosh \eta = \sinh \eta \tag{140}$$

Using Eq.(139) and Eq.(140), the Lorentz transformation (135) can be written as:

$$\begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \cosh \eta & \sinh \eta & 0 & 0 \\ \sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^0 \cosh \eta + x^1 \sinh \eta \\ x^0 \sinh \eta + x^1 \cosh \eta \\ x^2 \\ x^3 \end{pmatrix} \tag{141}$$

Eq.(141) suggests that the Lorentz transformation is like a rotation with hyperbolic functions.

1.11.6 Transformation of ψ due to Spatial Rotation of x

We want to examine how $\psi(x)$ changes due to each infinitesimal rotation $R_1(\delta\theta)$, $R_2(\delta\theta)$, and $R_3(\delta\theta)$ on x. However, there is one thing that I should tell you; we shall adopt $R_i(-\delta\theta)$, rather than $R_i(\delta\theta)$, if we want to agree with the mark and board analogy (Eq.(130) and Eq.(131)). Here is why: applying $R_i(\delta\theta)$ *on* x means rotating the board (or the marks) by $\delta\theta$, which is equivalent to rotating the arrows on the board by $-\delta\theta$. Then, to rotate the arrows by $\delta\theta$, we need to rotate the board by $-\delta\theta$. That's where the minus sign comes from. Hence, all we have to modify is replace $\delta\theta$ with $-\delta\theta$ in Eq.(134).

1) Infinitesimal change of $\psi(x)$ due to $R_1(-\delta\theta)$ on x

$$\begin{aligned}
\psi(x') =& \psi(x^0,\, x^1,\, x^2 + x^3\delta\theta,\, -x^2\delta\theta + x^3) \\
=& \psi(x^0,\, x^1,\, x^2,\, x^3) + x^3\delta\theta\frac{\partial\psi}{\partial x^2} - x^2\delta\theta\frac{\partial\psi}{\partial x^3} \\
=& \left(1 + x^3\delta\theta\frac{\partial}{\partial x^2} - x^2\delta\theta\frac{\partial}{\partial x^3}\right)\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(x^3\left(-i\frac{\partial}{\partial x^2}\right) - x^2\left(-i\frac{\partial}{\partial x^3}\right)\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(x^3 p_2 - x^2 p_3\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left(1 - i\delta\theta L_1\right)\psi(x^0,\, x^1,\, x^2,\, x^3)
\end{aligned} \tag{142}$$

where $L_1 = x^2 p_3 - x^3 p_2 = -i\left(x^2 \frac{\partial}{\partial x^3} - x^3 \frac{\partial}{\partial x^2}\right)$

2) Infinitesimal change of $\psi(x)$ due to $R_2(-\delta\theta)$ on x

$$
\begin{aligned}
\psi(x') =& \psi(x^0,\, x^1 - x^3\delta\theta,\, x^2,\, x^1\delta\theta + x^3) \\
=& \psi(x^0,\, x^1,\, x^2,\, x^3) - x^3\delta\theta\frac{\partial\psi}{\partial x^1} + x^1\delta\theta\frac{\partial\psi}{\partial x^3} \\
=& \left(1 - x^3\delta\theta\frac{\partial}{\partial x^1} + x^1\delta\theta\frac{\partial}{\partial x^3}\right)\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(-x^3\left(-i\frac{\partial}{\partial x^1}\right) + x^1\left(-i\frac{\partial}{\partial x^3}\right)\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(-x^3 p_1 + x^1 p_3\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left(1 - i\delta\theta L_2\right)\psi(x^0,\, x^1,\, x^2,\, x^3)
\end{aligned}
\tag{143}
$$

where $L_2 = x^3 p_1 - x^1 p_3 = -i\left(x^3 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^3}\right)$

3) Infinitesimal change of $\psi(x)$ due to $R_3(-\delta\theta)$

$$
\begin{aligned}
\psi(x') =& \psi(x^0,\, x^1 + x^2\delta\theta,\, -x^1\delta\theta + x^2,\, x^3) \\
=& \psi(x^0,\, x^1,\, x^2,\, x^3) + x^2\delta\theta\frac{\partial\psi}{\partial x^1} - x^1\delta\theta\frac{\partial\psi}{\partial x^2} \\
=& \left(1 + x^2\delta\theta\frac{\partial}{\partial x^1} - x^1\delta\theta\frac{\partial}{\partial x^2}\right)\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(x^2\left(-i\frac{\partial}{\partial x^1}\right) - x^1\left(-i\frac{\partial}{\partial x^2}\right)\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left[1 - \frac{1}{i}\delta\theta\left(x^2 p_1 - x^1 p_2\right)\right]\psi(x^0,\, x^1,\, x^2,\, x^3) \\
=& \left(1 - i\delta\theta L_3\right)\psi(x^0,\, x^1,\, x^2,\, x^3)
\end{aligned}
\tag{144}
$$

where $L_3 = x^1 p_2 - x^2 p_1 = -i\left(x^1 \frac{\partial}{\partial x^2} - x^2 \frac{\partial}{\partial x^1}\right)$

Suppose we successively take the infinitesimal spatial rotation in each axis:

$$
\begin{aligned}
&\left(1 - i\delta\theta_i L_i\right)\left(1 - i\delta\theta_j L_j\right)\left(1 - i\delta\theta L_k\right)\psi(x) \\
=&(1 - i\delta\theta_i L_i - i\delta\theta_j L_j - i\delta\theta_k L_k)\psi(x) \\
=&(1 - i\vec{L}\cdot\delta\vec{\theta})\psi(x)
\end{aligned}
\tag{145}
$$

up to the first order of $\delta\theta_a$ $(a = i, j, k)$.

Eq.(145) describes how $\psi(x)$ infinitesimally changes for a given infinitesimal rotation of $\psi(x)$. Then how about a finite rotation by $\Delta\vec{\theta}$? We apply the infinitesimal operator many times!

$$
\begin{aligned}
&\lim_{N\to\infty}(1-i\vec{L}\cdot\delta\vec{\theta})^N\psi(x) = \lim_{N\to\infty}\left(1-i\vec{L}\cdot\frac{\Delta\vec{\theta}}{N}\right)^N\psi(x)\\
=&\lim_{N\to\infty}\left(1-i\vec{L}\cdot\frac{\Delta\vec{\theta}}{N}\right)^{\frac{N}{-i\vec{L}\cdot\Delta\vec{\theta}}\cdot\frac{-i\vec{L}\cdot\Delta\vec{\theta}}{N}N}\psi(x) \qquad (146)\\
=&\left[\lim_{N\to\infty}\left(1-i\vec{L}\cdot\frac{\Delta\vec{\theta}}{N}\right)^{\frac{N}{-i\vec{L}\cdot\Delta\vec{\theta}}}\right]^{-i\vec{L}\cdot\Delta\vec{\theta}}\psi(x)
\end{aligned}
$$

What is inside the square brackets above is simply e. Thus,

$$\psi'(x') = e^{-i\vec{L}\cdot\Delta\vec{\theta}}\psi(x) \tag{147}$$

1.11.7 Transformation of ψ due to the Lorentz Transformation on x

Likewise, we take η to $-\eta$ in Eq.(141). The Taylor expansions of $\cosh\eta$ and $\sinh\eta$ are:

$$
\begin{aligned}
\cosh\eta &= 1+\frac{\eta^2}{2!}+\frac{\eta^4}{4!}+\frac{\eta^6}{6!}+\cdots\\
\sinh\eta &= \frac{\eta^1}{1!}+\frac{\eta^3}{3!}+\frac{\eta^5}{5!}+\frac{\eta^7}{7!}+\cdots
\end{aligned} \tag{148}
$$

Thus, for a small $\eta\to d\eta$,

$$\cosh d\eta \approx 1 \qquad\qquad \sinh d\eta \approx d\eta \tag{149}$$

Then Eq.(141), with $\eta\to-\eta$, becomes:

$$
\begin{pmatrix} x^{0'}\\ x^{1'}\\ x^{2'}\\ x^{3'}\end{pmatrix} = \begin{pmatrix} 1 & -d\eta & 0 & 0\\ -d\eta & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} x^0\\ x^1\\ x^2\\ x^3\end{pmatrix} = \begin{pmatrix} x^0 - x^1 d\eta\\ -x^0 d\eta + x^1\\ x^2\\ x^3\end{pmatrix} \tag{150}
$$

Let us consider the infinitesimal change of $\psi(x)$ due to the infinitesimal Lorentz transformation.

$$
\begin{aligned}
\psi'(x') =& \psi(x^0 - x^1 d\eta,\ -x^0 d\eta + x^1,\ x^2,\ x^3)\\
=&\psi(x^0,\ x^1,\ x^2,\ x^3) - x^1 d\eta\frac{\partial\psi}{\partial x^0} - x^0 d\eta\frac{\partial\psi}{\partial x^1}\\
=&\left(1 - x^1 d\eta\frac{\partial}{\partial x^0} - x^0 d\eta\frac{\partial}{\partial x^1}\right)\psi\\
=&\left[1-\frac{1}{i}\left(-x^1 d\eta\left(-i\frac{\partial}{\partial x^0}\right) - x^0 d\eta\left(-i\frac{\partial}{\partial x^1}\right)\right)\right]\psi\\
=&\left[1 - id\eta\left(-ix^1\frac{\partial}{\partial x^0} - ix^0\frac{\partial}{\partial x^1}\right)\right]\psi\\
=&\left[1 - id\eta N_1\right]\psi
\end{aligned} \tag{151}
$$

where $N_1 = -i\left(x^1 \frac{\partial}{\partial x^0} + x^0 \frac{\partial}{\partial x^1}\right)$

Eq.(151) considers the infinitesimal Lorentz transformation into the x-axis. Because there is nothing special about the x-axis over the y and z-axis, N_2 and N_3 can be obtained by following:

$$\begin{aligned} x^1 \to x^2 : \qquad & N_2 = -i\left(x^2 \frac{\partial}{\partial x^0} + x^0 \frac{\partial}{\partial x^2}\right) \\ x^1 \to x^3 : \qquad & N_3 = -i\left(x^3 \frac{\partial}{\partial x^0} + x^0 \frac{\partial}{\partial x^3}\right) \end{aligned} \tag{152}$$

which correspond to the infinitesimal Lorentz transformation along the y and z-axis, respectively.

Suppose we successively take the infinitesimal Lorentz transformation along each axis:

$$\begin{aligned} &(1 - i\delta\eta_i N_i)(1 - i\delta\eta_j N_j)(1 - i\delta\eta_k N_k)\psi(x) \\ =&(1 - i\delta\eta_i N_i - i\delta\eta_j N_j - i\delta\eta_k N_k)\psi(x) \\ =&(1 - i\vec{N} \cdot \delta\vec{\eta})\psi(x) \end{aligned} \tag{153}$$

up to the first order of $\delta\eta_a$ $(a = i, j, k)$.

Eq.(153) describes how $\psi(x)$ infinitesimally changes for a given infinitesimal Lorentz transformation of $\psi(x)$. Then how about a finite Lorentz transformation by $\Delta\vec{\eta}$? Again, we apply the infinitesimal operator many times!

$$\begin{aligned} \lim_{M\to\infty}(1 - i\vec{N} \cdot \delta\vec{\eta})^M \psi(x) &= \lim_{M\to\infty}\left(1 - i\vec{N} \cdot \frac{\Delta\vec{\eta}}{M}\right)^M \psi(x) \\ = \lim_{M\to\infty}\left(1 - i\vec{N} \cdot \frac{\Delta\vec{\eta}}{M}\right)^{\frac{M}{-i\vec{N}\cdot\Delta\vec{\eta}} \cdot \frac{-i\vec{N}\cdot\Delta\vec{\eta}}{M} M} &\psi(x) \\ = \left[\lim_{M\to\infty}\left(1 - i\vec{N} \cdot \frac{\Delta\vec{\eta}}{M}\right)^{\frac{M}{-i\vec{N}\cdot\Delta\vec{\eta}}}\right]^{-i\vec{N}\cdot\Delta\vec{\eta}} &\psi(x) \end{aligned} \tag{154}$$

What is inside the square brackets above is simply e. Thus,

$$\psi'(x') = e^{-i\vec{N}\cdot\Delta\vec{\eta}}\psi(x) \tag{155}$$

1.11.8 Combining the Rotation and Lorentz Transformation

Suppose we take both the rotation and Lorentz transformation on $\psi(x)$. Then, by Eq.(147) and Eq.(155),

$$\begin{aligned} \psi'(x') &= e^{-i\vec{N}\cdot\Delta\vec{\eta}} e^{-i\vec{L}\cdot\Delta\vec{\theta}}\psi(x) \\ &= e^{-i(\vec{N}\cdot\Delta\vec{\eta} + \vec{L}\cdot\Delta\vec{\theta})}\psi(x) \\ &= e^{-i(N_1\eta_1 + N_2\eta_2 + N_3\eta_3 + L_1\theta_1 + L_2\theta_2 + L_3\theta_3)}\psi(x) \end{aligned} \tag{156}$$

It would be nice if there is a way to write down the last line of Eq.(156) more neatly. By writing $\frac{\partial}{\partial x^\mu} = \partial_\mu$, L_i and N_i can be written as:

$$L_i = -i(x^j\partial_k - x^k\partial_j) \qquad\qquad N_i = -i(x^i\partial_0 + x^0\partial_i) \tag{157}$$

By using $\partial^0 = \partial_0$ and $\partial^k = -\partial_k$,

$$L_i = i(x^j\partial^k - x^k\partial^j) \qquad\qquad N_i = i(x^0\partial^i - x^i\partial^0) \tag{158}$$

To write Eq.(158) all together, let us define the Lorentz transformation generator $M^{\mu\nu}$ and its variables $\omega_{\mu\nu}$ by:

$$M^{\mu\nu} = i(x^\mu\partial^\nu - x^\nu\partial^\mu) \tag{159}$$

$$\begin{aligned}(M^{23}, M^{31}, M^{12}) &= (L_1, L_2, L_3)\\ (M^{01}, M^{02}, M^{03}) &= (N_1, N_2, N_3)\end{aligned} \tag{160}$$

$$\begin{aligned}(\omega_{23}, \omega_{31}, \omega_{12}) &= (\theta_1, \theta_2, \theta_3)\\ (\omega_{01}, \omega_{02}, \omega_{03}) &= (\eta_1, \eta_2, \eta_3)\end{aligned} \tag{161}$$

where the diagonal terms of $M^{\mu\nu}$ and $\omega_{\mu\nu}$ are zero:

$$M^{\alpha\alpha} = 0 \qquad\qquad \omega_{\alpha\alpha} = 0 \tag{162}$$

Notice that $M^{\mu\nu}$ *must* be antisymmetric (due to Eq.(159)) and we *can* define $\omega_{\mu\nu}$ to be antisymmetric:

$$M^{\nu\mu} = -M^{\mu\nu} \qquad\qquad \omega_{\nu\mu} = -\omega_{\mu\nu} \tag{163}$$

Then the exponent of the last line of Eq.(156) is written as:

$$\begin{aligned}
&N_1\eta_1 + N_2\eta_2 + N_3\eta_3 + L_1\theta_1 + L_2\theta_2 + L_3\theta_3\\
&= M^{01}\omega_{01} + M^{02}\omega_{02} + M^{03}\omega_{03} + M^{23}\omega_{23} + M^{31}\omega_{31} + M^{12}\omega_{12}\\
&= \frac{1}{2}(M^{01}\omega_{01} + M^{02}\omega_{02} + M^{03}\omega_{03} + M^{23}\omega_{23} + M^{31}\omega_{31} + M^{12}\omega_{12}\\
&+ M^{10}\omega_{10} + M^{20}\omega_{20} + M^{30}\omega_{30} + M^{32}\omega_{32} + M^{13}\omega_{13} + M^{21}\omega_{21}\\
&+ M^{00}\omega_{00} + M^{11}\omega_{11} + M^{22}\omega_{22} + M^{33}\omega_{33})\\
&= \frac{1}{2}M^{\mu\nu}\omega_{\mu\nu}
\end{aligned} \tag{164}$$

where we used

$$\begin{aligned}&M^{01}\omega_{01} + M^{02}\omega_{02} + M^{03}\omega_{03} + M^{23}\omega_{23} + M^{31}\omega_{31} + M^{12}\omega_{12}\\ =&M^{10}\omega_{10} + M^{20}\omega_{20} + M^{30}\omega_{30} + M^{32}\omega_{32} + M^{13}\omega_{13} + M^{21}\omega_{21}\end{aligned} \tag{165}$$

because $M^{\mu\nu}$ and $\omega_{\mu\nu}$ are *both* antisymmetric (therefore $M^{\mu\nu}\omega_{\mu\nu} = M^{\nu\mu}\omega_{\nu\mu}$) and any term involving either $M^{\alpha\alpha}$ or $\omega_{\alpha\alpha}$ is zero.

Therefore, with Eq.(164), Eq.(156) can be written as:

$$\psi'(x') = e^{-\frac{i}{2}M^{\mu\nu}\omega_{\mu\nu}}\psi(x) \tag{166}$$

1.12 Is the Dirac Equation invariant under the Lorentz transformation?

The laws of physics must be the same in all inertial frames. Is that going to be the case for the Dirac equation too? We will see.

1.12.1 A Modification on Transformation of $\psi(x)$

So we have discovered how $\psi(x)$ transforms. Let us go back to Eq.(164) again, which is a somewhat crude form:

$$N_1\eta_1 + N_2\eta_2 + N_3\eta_3 + L_1\theta_1 + L_2\theta_2 + L_3\theta_3 = \frac{1}{2}M^{\mu\nu}\omega_{\mu\nu} \tag{167}$$

Here $L_i = i(x^j\partial^k - x^k\partial^j)$ is an orbital angular momentum operator. However, we know, from quantum mechanics, that orbital angular momentum operators have the same commutation relation with spin angular momentum operators. So why not replacing L_i with the four-dimensional extension of spin operator, $\frac{1}{2}\Sigma_i$?

$$L_i \to \frac{1}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} = \frac{1}{2}\Sigma_i \tag{168}$$

And some clever people found that the spatial components of $S^{\mu\nu}$, defined by the following, are given by $\frac{1}{2}\Sigma_i$:

$$S^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right] \qquad\qquad S^{\mu\nu} = -S^{\nu\mu} \tag{169}$$

Let us use Eq.(106) and Eq.(110), which are

$$\{\gamma_i,\, \gamma_j\} = 0 \quad (i \neq j) \qquad\qquad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}$$

to compute Eq.(169). Then the spatial components, S^{ij}, are given by:

$$\begin{aligned} S^{ij} &= \frac{i}{4}\left[\gamma^i, \gamma^j\right] = \frac{i}{4}\left(\gamma^i\gamma^j - \gamma^j\gamma^i\right) = \frac{i}{2}\gamma^i\gamma^j \\ &= \frac{i}{4}\begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}\begin{pmatrix} 0 & -\sigma_j \\ \sigma_j & 0 \end{pmatrix} \\ &= -\frac{i}{4}\begin{pmatrix} \sigma_i\sigma_j & 0 \\ 0 & \sigma_i\sigma_j \end{pmatrix} \\ &= -\frac{i}{4}\begin{pmatrix} 2i\epsilon_{ijk}\sigma_k & 0 \\ 0 & 2i\epsilon_{ijk}\sigma_k \end{pmatrix} \\ &= \frac{1}{2}\epsilon_{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \end{aligned} \tag{170}$$

Therefore,

$$S^{ij} = \frac{1}{2}\epsilon_{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} = \frac{1}{2}\Sigma_k \quad (i \neq j) \tag{171}$$

as we wanted in Eq.(168). Thus, we can say L_i in Eq.(167) can be replaced by Eq.(171). In other words, M^{ij} is replaced by S^{ij} (See Eq.(160)). Then we *must* accept that M^{0i} is replaced by S^{0i}. In other words, N_i is replaced by $\frac{i}{2}\alpha_i$.

$$\begin{aligned} S^{0i} &= \frac{i}{4}\left[\gamma^0, \gamma^i\right] = \frac{i}{4}\left(\gamma^0\gamma^i - \gamma^i\gamma^0\right) = \frac{i}{4}\left(\beta\beta\alpha_i - \beta\alpha_i\beta\right) \\ &= \frac{i}{4}\left(\beta\beta\alpha_i + \beta\beta\alpha_i\right) = \frac{i}{4}\left(\alpha_i + \alpha_i\right) = \frac{i}{2}\alpha_i \end{aligned} \tag{172}$$

Therefore,

$$S^{0i} = \frac{i}{2}\begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \tag{173}$$

Therefore, Eq.(167) changes into the following:

$$\frac{1}{2}M^{\mu\nu}\omega_{\mu\nu} \to \frac{1}{2}S^{\mu\nu}\omega_{\mu\nu} \tag{174}$$

Then Eq.(166) changes into the following:

$$\psi'(x') = e^{-\frac{i}{2}S^{\mu\nu}\omega_{\mu\nu}}\psi(x) \tag{175}$$

Let us calculate the explicit form of $-\frac{i}{2}S^{\mu\nu}\omega_{\mu\nu}$. To do that, we follow a similar calculation as we did in Eq.(164). Keeping in mind that

$$N_i \to \frac{i}{2}\alpha_i \qquad\qquad L_i \to \frac{1}{2}\Sigma_i \tag{176}$$

$$\begin{aligned} &\frac{i}{2}\alpha_1\eta_1 + \frac{i}{2}\alpha_2\eta_2 + \frac{i}{2}\alpha_3\eta_3 + \frac{1}{2}\Sigma_1\theta_1 + \frac{1}{2}\Sigma_2\theta_2 + \frac{1}{2}\Sigma_3\theta_3 \\ &= S^{01}\omega_{01} + S^{02}\omega_{02} + S^{03}\omega_{03} + S^{23}\omega_{23} + S^{31}\omega_{31} + S^{12}\omega_{12} \\ &= \frac{1}{2}(S^{01}\omega_{01} + S^{02}\omega_{02} + S^{03}\omega_{03} + S^{23}\omega_{23} + S^{31}\omega_{31} + S^{12}\omega_{12} \\ &+ S^{10}\omega_{10} + S^{20}\omega_{20} + S^{30}\omega_{30} + S^{32}\omega_{32} + S^{13}\omega_{13} + S^{21}\omega_{21} \\ &+ S^{00}\omega_{00} + S^{11}\omega_{11} + S^{22}\omega_{22} + S^{33}\omega_{33}) \\ &= \frac{1}{2}S^{\mu\nu}\omega_{\mu\nu} \end{aligned} \tag{177}$$

where we used

$$\begin{aligned} &S^{01}\omega_{01} + S^{02}\omega_{02} + S^{03}\omega_{03} + S^{23}\omega_{23} + S^{31}\omega_{31} + S^{12}\omega_{12} \\ =&S^{10}\omega_{10} + S^{20}\omega_{20} + S^{30}\omega_{30} + S^{32}\omega_{32} + S^{13}\omega_{13} + S^{21}\omega_{21} \end{aligned} \tag{178}$$

because $S^{\mu\nu}$ and $\omega_{\mu\nu}$ are *both* antisymmetric (therefore $S^{\mu\nu}\omega_{\mu\nu} = S^{\nu\mu}\omega_{\nu\mu}$) and any term involving either $S^{\alpha\alpha}$ or $\omega_{\alpha\alpha}$ is zero.
Then, by Eq.(177),

$$\begin{aligned} -\frac{i}{2}S^{\mu\nu}\omega_{\mu\nu} &= \frac{1}{2}\left(\alpha_1\eta_1 + \alpha_2\eta_2 + \alpha_3\eta_3\right) - \frac{i}{2}\left(\Sigma_1\theta_1 + \Sigma_2\theta_2 + \Sigma_3\theta_3\right) \\ &= \frac{1}{2}\vec{\alpha}\cdot\vec{\eta} - \frac{i}{2}\vec{\Sigma}\cdot\vec{\theta} \end{aligned} \tag{179}$$

Thus the transformation rule in Eq.(175) can be written as:

$$\psi'(x') = e^{-\frac{i}{2}S^{\mu\nu}\omega_{\mu\nu}}\psi(x) = e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}-\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}}\psi(x) \tag{180}$$

So far, everything is the same except we replaced $M_{\mu\nu}$ to $S_{\mu\nu}$. Let us define

$$S(L) = e^{-\frac{i}{2}S^{\mu\nu}\omega_{\mu\nu}} = e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}-\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}} \tag{181}$$

The inverse of $S(L)$ can be obtained by "unrotating" $\psi(x)$. That is, $\omega_{\mu\nu} \to -\omega_{\mu\nu}$. Let us simply our notation a bit ($S(L) = S$).

$$S^{-1} = e^{-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}} \tag{182}$$

Its Hermitian conjugate is given by:

$$S^{\dagger} = e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}} \tag{183}$$

because α_i and Σ_i are Hermitian, because the Pauli matrices are Hermitian (Eq.(43) and Eq.(171)). Eq.(182) and Eq.(183) are related as follows.

$$S^{-1} = \gamma^0 S^{\dagger}\gamma^0 \tag{184}$$

Here's proof.

$$\begin{aligned}\gamma^0 S^{\dagger}\gamma^0 &= \gamma^0 e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}}\gamma^0 \\ &= \gamma^0\left[1 + \frac{1}{1!}\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta} + \frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right) + \cdots\right]\gamma^0\end{aligned} \tag{185}$$

However, we know

$$\begin{aligned}\gamma^0\alpha_i\gamma^0 &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}0 & \sigma_i\\ -\sigma_i & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \\ &= \begin{pmatrix}\sigma_i & 0\\ 0 & -\sigma_i\end{pmatrix} = -\begin{pmatrix}-\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} = -\alpha_i\end{aligned} \tag{186}$$

$$\begin{aligned}\gamma^0\Sigma_i\gamma^0 &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} = \Sigma_i\end{aligned} \tag{187}$$

Thus the second term in Eq.(185) is:

$$\begin{aligned}\frac{1}{1!}\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta} + \frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\gamma^0 &= \frac{1}{1!}\left(\frac{1}{2}\gamma^0\vec{\alpha}\gamma^0\cdot\vec{\eta} + \frac{i}{2}\gamma^0\vec{\Sigma}\gamma^0\cdot\vec{\theta}\right) \\ &= \frac{1}{1!}\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta} + \frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\end{aligned} \tag{188}$$

How about other terms in Eq.(185)? We can use $(\gamma^0)^2 = I$. For example,

$$
\begin{aligned}
&\frac{1}{2!}\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)^2\gamma^0 \\
&=\frac{1}{2!}\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\gamma^0 \\
&=\frac{1}{2!}\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\gamma^0\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\gamma^0 \\
&=\frac{1}{2!}\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right) \\
&=\frac{1}{2!}\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)^2
\end{aligned}
\tag{189}
$$

Thus we can say:

$$
\frac{1}{n!}\gamma^0\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)^n\gamma^0=\frac{1}{n!}\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)^n
\tag{190}
$$

Going back to Eq.(185),

$$
\begin{aligned}
\gamma^0 S^\dagger\gamma^0 &= \gamma^0 e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}}\gamma^0 \\
&=\gamma^0\left[1+\frac{1}{1!}\left(\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)+\cdots\right]\gamma^0 \\
&=1+\frac{1}{1!}\left(-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}\right)+\cdots \\
&=e^{-\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}+\frac{i}{2}\vec{\Sigma}\cdot\vec{\theta}}=S^{-1}
\end{aligned}
\tag{191}
$$

Thus, identity (184) is proven. Then we have the following relations:

$$
\begin{aligned}
&\psi(x)\to\psi'(x')=S\psi(x) \\
&\bar{\psi}(x)\to\bar{\psi}'(x')=[\psi'(x')]^\dagger\gamma^0=[S\psi(x)]^\dagger\gamma^0=\psi^\dagger(x)S^\dagger\gamma^0 \\
&\qquad\qquad\quad=\psi^\dagger(x)\gamma^0\gamma^0S^\dagger\gamma^0=\bar{\psi}(x)S^{-1}
\end{aligned}
\tag{192}
$$

Now let me provide a mathematical justification on how $\bar{\psi}(x)\gamma^\mu\psi(x)$ transforms like a 4-vector (See Eq.(121)). For simplicity, let us consider only the Lorentz boost in the x direction. That is,

$$
\vec{\eta}=(\eta_1,0,0) \qquad\qquad \vec{\theta}=(0,0,0)
$$

Then, from Eq.(181), $S(L)$ is written as (Keep in mind that $\alpha_i^2 = 1$, which was one of the requirements of the Dirac equation):

$$
\begin{aligned}
S(L) &= e^{\frac{1}{2}\alpha_1\eta_1} \\
&= 1 + \frac{1}{1!}\left(\frac{1}{2}\alpha_1\eta_1\right)^1 + \frac{1}{2!}\left(\frac{1}{2}\alpha_1\eta_1\right)^2 + \frac{1}{3!}\left(\frac{1}{2}\alpha_1\eta_1\right)^3 + \frac{1}{4!}\left(\frac{1}{2}\alpha_1\eta_1\right)^4 + \cdots \\
&= 1 + \alpha_1\frac{1}{1!}\left(\frac{1}{2}\eta_1\right)^1 + \alpha_1^2\frac{1}{2!}\left(\frac{1}{2}\eta_1\right)^2 + \alpha_1^3\frac{1}{3!}\left(\frac{1}{2}\eta_1\right)^3 + \alpha_1^4\frac{1}{4!}\left(\frac{1}{2}\eta_1\right)^4 + \cdots \\
&= 1 + \alpha_1\frac{1}{1!}\left(\frac{1}{2}\eta_1\right)^1 + \frac{1}{2!}\left(\frac{1}{2}\eta_1\right)^2 + \alpha_1\frac{1}{3!}\left(\frac{1}{2}\eta_1\right)^3 + \frac{1}{4!}\left(\frac{1}{2}\eta_1\right)^4 + \cdots \\
&= \left[1 + \frac{1}{2!}\left(\frac{1}{2}\eta_1\right)^2 + \frac{1}{4!}\left(\frac{1}{2}\eta_1\right)^4 + \cdots\right] + \alpha_1\left[\frac{1}{1!}\left(\frac{1}{2}\eta_1\right)^1 + \frac{1}{3!}\left(\frac{1}{2}\eta_1\right)^3 + \cdots\right] \\
&= \cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}
\end{aligned}
\tag{193}
$$

1.12.2 $\bar{\psi}\gamma^\mu\psi$ is a 4-vector

Then, due to Eq.(192), the Lorentz-boosted $\bar{\psi}\gamma^\mu\psi$ can be written as:

$$
\begin{aligned}
\bar{\psi}'\gamma^\mu\psi' &= \bar{\psi}S^{-1}\gamma^\mu S\psi = (\psi^\dagger\gamma^0)(\gamma^0 S^\dagger\gamma^0)\gamma^\mu S\psi \\
&= \psi^\dagger S^\dagger\gamma^0\gamma^\mu S\psi
\end{aligned}
\tag{194}
$$

If $\mu = 0$,

$$
\begin{aligned}
\bar{\psi}'\gamma^0\psi' &= \psi^\dagger S^\dagger\gamma^0\gamma^0 S\psi = \psi^\dagger S^\dagger S\psi \\
&= \psi^\dagger\left[\cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}\right]\left[\cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\cosh^2\frac{\eta_1}{2} + \sinh^2\frac{\eta_1}{2} + 2\alpha_1\sinh\frac{\eta_1}{2}\cosh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\cosh\eta_1 + \alpha_1\sinh\eta_1\right]\psi \\
&= \psi^\dagger\gamma^0\gamma^0\left[\cosh\eta_1 + \alpha_1\sinh\eta_1\right]\psi \\
&= \bar{\psi}\gamma^0\cosh\eta_1\psi + \bar{\psi}\gamma^1\sinh\eta_1\psi \\
&= \bar{\psi}\gamma^0\gamma\psi + \bar{\psi}\gamma^1\beta\gamma\psi \\
&= \gamma\bar{\psi}\gamma^0\psi + \beta\gamma\bar{\psi}\gamma^1\psi
\end{aligned}
\tag{195}
$$

where we used

$$\alpha_1^\dagger = \alpha_1 \qquad\qquad \gamma^0\alpha_1 = \gamma^1$$

$$\cosh^2 x + \sinh^2 x = \cosh 2x \qquad\qquad 2\sinh x\cosh x = \sinh 2x$$

If $\mu = 1$,

$$
\begin{aligned}
\bar{\psi}'\gamma^1\psi' &= \psi^\dagger S^\dagger \gamma^0 \gamma^1 S\psi = \psi^\dagger S^\dagger \gamma^0\gamma^0\alpha_1 S\psi \\
&= \psi^\dagger \left[\cosh\frac{\eta_1}{2} + \alpha_1 \sinh\frac{\eta_1}{2}\right]\alpha_1\left[\cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\alpha_1\cosh^2\frac{\eta_1}{2} + \alpha_1\sinh^2\frac{\eta_1}{2} + 2\sinh\frac{\eta_1}{2}\cosh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\alpha_1\cosh\eta_1 + \sinh\eta_1\right]\psi \\
&= \psi^\dagger\gamma^0\gamma^0\left[\alpha_1\cosh\eta_1 + \sinh\eta_1\right]\psi \\
&= \bar{\psi}\gamma^1\cosh\eta_1\psi + \bar{\psi}\gamma^0\sinh\eta_1\psi \\
&= \bar{\psi}\gamma^1\gamma\psi + \bar{\psi}\gamma^0\beta\gamma\psi \\
&= \gamma\bar{\psi}\gamma^1\psi + \beta\gamma\bar{\psi}\gamma^0\psi
\end{aligned}
\tag{196}
$$

If $\mu = i$ (where $i = 2, 3$),

$$
\begin{aligned}
\bar{\psi}'\gamma^i\psi' &= \psi^\dagger S^\dagger\gamma^0\gamma^i S\psi = \psi^\dagger S^\dagger\gamma^0\gamma^0\alpha_i S\psi \\
&= \psi^\dagger\left[\cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}\right]\alpha_i\left[\cosh\frac{\eta_1}{2} + \alpha_1\sinh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\alpha_i\cosh^2\frac{\eta_1}{2} + \alpha_1\alpha_i\alpha_1\sinh^2\frac{\eta_1}{2} + \{\alpha_1,\alpha_i\}\sinh\frac{\eta_1}{2}\cosh\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\left[\alpha_i\cosh^2\frac{\eta_1}{2} - \alpha_i\sinh^2\frac{\eta_1}{2}\right]\psi \\
&= \psi^\dagger\gamma^0\gamma^0\alpha_i\psi \\
&= \bar{\psi}\gamma^i\psi
\end{aligned}
\tag{197}
$$

where we used Eq.(29(ii)).
Restating Eq.(195), (196), and (197),

$$
\begin{aligned}
\bar{\psi}'\gamma^0\psi' &= \gamma\bar{\psi}\gamma^0\psi + \beta\gamma\bar{\psi}\gamma^1\psi \\
\bar{\psi}'\gamma^1\psi' &= \gamma\bar{\psi}\gamma^1\psi + \beta\gamma\bar{\psi}\gamma^0\psi \\
\bar{\psi}'\gamma^i\psi' &= \bar{\psi}\gamma^i\psi \qquad (i = 2, 3)
\end{aligned}
\tag{198}
$$

By comparing Eq.(198) and Eq.(135), we see that $\bar{\psi}\gamma^\mu\psi$ obeys the Lorentz transformation. Or we can simply write:

$$\bar{\psi}'\gamma^\mu\psi' = \Lambda^\mu_\nu\bar{\psi}\gamma^\nu\psi \tag{199}$$

where Λ^μ_ν is the Lorentz transformation matrix element in Eq.(135). From Eq.(194) and Eq.(199), we find the following relation:

$$\bar{\psi}'\gamma^\mu\psi' = \bar{\psi}S^{-1}\gamma^{\mu'}S\psi = \Lambda^\mu_\nu\bar{\psi}\gamma^\nu\psi \tag{200}$$

By omitting $\psi(x)$ and $\bar{\psi}(x)$,

$$S^{-1}\gamma^\mu S = \Lambda^\mu_\nu\gamma^\nu \tag{201}$$

1.12.3 Lorentz Invariance of the Dirac Equation

Let us finally verify if the Dirac equation is invariant under the Lorentz transformation. So we start from the Lorentz-transformed Dirac equation and see if we can end up with its original form.

$$\begin{aligned}
&(\gamma^{\mu'} i\partial_{\mu'} - m)\psi'(x') = 0 \\
&(\gamma^{\mu'} i\Lambda^{\nu}_{\mu'}\partial_{\nu} - m)S\psi(x) = 0 \\
&S^{-1}(\gamma^{\mu'} i\Lambda^{\nu}_{\mu'}\partial_{\nu} - m)S\psi(x) = 0 \\
&(S^{-1}\gamma^{\mu'} S i\Lambda^{\nu}_{\mu'}\partial_{\nu} - m)\psi(x) = 0 \\
&(\Lambda^{\mu'}_{\sigma}\gamma^{\sigma} i\Lambda^{\nu}_{\mu'}\partial_{\nu} - m)\psi(x) = 0 \\
&(i\delta^{\nu}_{\sigma}\gamma^{\sigma}\partial_{\nu} - m)\psi(x) = 0 \\
&(i\gamma^{\nu}\partial_{\nu} - m)\psi(x) = 0
\end{aligned} \tag{202}$$

where we used Eq.(201) on going from the fourth to the fifth line. $\partial_{\mu'}$ means partial derivative in x' coordinate system. Therefore, the Dirac equation is invariant under the Lorentz transformation, respecting the special theory of relativity.

1.13 Plane-Wave Solution

Let us solve the Dirac equation:

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

By multiplying γ^0 from the left and using

$$\gamma^0 = \beta \qquad \gamma^i = \gamma^0\alpha_i \qquad (\gamma^0)^2 = I$$

The Dirac equation becomes:

$$\begin{aligned}
&\gamma^0(i\gamma^{\mu}\partial_{\mu} - m)\psi(x) = 0 \\
&\gamma^0(i\gamma^0\partial_0 + i\gamma^k\partial_k - m)\psi(x) = 0 \\
&\gamma^0(i\gamma^0\partial_0 + i\gamma^0\sum_{k=1}^{3}\alpha_k\partial_k - m)\psi(x) = 0 \\
&(i\partial_0 + i\sum_{k=1}^{3}\alpha_k\partial_k - m\gamma^0)\psi(x) = 0
\end{aligned} \tag{203}$$

Keeping in mind that

$$\alpha_k = \begin{pmatrix} -\sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \qquad \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The last line of Eq.(203) becomes

$$
\begin{aligned}
&\left[\begin{pmatrix} i\partial_0 & 0 \\ 0 & i\partial_0 \end{pmatrix} + i\sum_{k=1}^{3}\begin{pmatrix} -\sigma_k\partial_k & 0 \\ 0 & \sigma_k\partial_k \end{pmatrix}\right]\begin{pmatrix}\phi_L(x) \\ \phi_R(x)\end{pmatrix} = \begin{pmatrix} 0 & m \\ m & 0\end{pmatrix}\begin{pmatrix}\phi_L(x) \\ \phi_R(x)\end{pmatrix} \\
&\begin{pmatrix} i\partial_0 - \vec{\sigma}\cdot i\nabla & 0 \\ 0 & i\partial_0 + \vec{\sigma}\cdot i\nabla \end{pmatrix}\begin{pmatrix}\phi_L(x) \\ \phi_R(x)\end{pmatrix} = \begin{pmatrix} m\phi_R(x) \\ m\phi_L(x)\end{pmatrix} \\
&\begin{pmatrix} p^0 + \vec{\sigma}\cdot\vec{p} & 0 \\ 0 & p^0 - \vec{\sigma}\cdot\vec{p} \end{pmatrix}\begin{pmatrix}\phi_L(x) \\ \phi_R(x)\end{pmatrix} = \begin{pmatrix} m\phi_R(x) \\ m\phi_L(x)\end{pmatrix}
\end{aligned}
\tag{204}
$$

where we used:

$$
-i\nabla = \vec{p} \qquad\qquad i\partial_0 = E = p^0 \tag{205}
$$

We can keep expanding Eq.(204):

$$
\begin{aligned}
(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(x) &= m\phi_R(x) \\
(p^0 - \vec{\sigma}\cdot\vec{p})\phi_R(x) &= m\phi_L(x)
\end{aligned}
\tag{206}
$$

Let us try a plane-wave solution:

$$
\phi_L(x) = N\phi_L(\vec{p})e^{-ip\cdot x} \qquad\qquad \phi_R(x) = N\phi_R(\vec{p})e^{-ip\cdot x} \tag{207}
$$

By plugging Eq.(207) into Eq.(206),

$$
\begin{aligned}
(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(\vec{p}) &= m\phi_R(\vec{p}) \\
(p^0 - \vec{\sigma}\cdot\vec{p})\phi_R(\vec{p}) &= m\phi_L(\vec{p})
\end{aligned}
\tag{208}
$$

We can use the first equation of Eq.(208) to eliminate $\phi_R(\vec{p})$ in the second equation of Eq.(208).

$$
\begin{aligned}
\phi_R(\vec{p}) = \frac{1}{m}(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(\vec{p}) &\xrightarrow{\text{Plug into}} (p^0 - \vec{\sigma}\cdot\vec{p})\phi_R(\vec{p}) = m\phi_L(\vec{p}) \\
(p^0 - \vec{\sigma}\cdot\vec{p})(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(\vec{p}) &= m^2\phi_L(\vec{p}) \\
\left[(p^0)^2 - (\vec{\sigma}\cdot\vec{p})^2\right]\phi_L(\vec{p}) &= m^2\phi_L(\vec{p})
\end{aligned}
\tag{209}
$$

Let us calculate $(\vec{\sigma}\cdot\vec{p})^2$:

$$
\begin{aligned}
(\vec{\sigma}\cdot\vec{p})^2 &= (\vec{\sigma}\cdot\vec{p})(\vec{\sigma}\cdot\vec{p}) \\
&= (\sigma_x p_x + \sigma_y p_y + \sigma_z p_z)(\sigma_x p_x + \sigma_y p_y + \sigma_z p_z) \\
&= (\sigma_x^2 p_x^2 + \sigma_x\sigma_y p_x p_y + \sigma_x\sigma_z p_x p_z \\
&+ \sigma_y\sigma_x p_y p_x + \sigma_y^2 p_y^2 + \sigma_y\sigma_z p_y p_z \\
&+ \sigma_z\sigma_x p_z p_x + \sigma_z\sigma_y p_z p_y + \sigma_z^2 p_z^2) \\
&= \left(p_x^2 + p_y^2 + p_z^2 + \{\sigma_x,\sigma_y\}p_x p_y + \{\sigma_x,\sigma_z\}p_x p_z + \{\sigma_y,\sigma_z\}p_y p_z\right) \\
&= p_x^2 + p_y^2 + p_z^2 \\
&= |\vec{p}|^2
\end{aligned}
\tag{210}
$$

where we used

$$\sigma_i^2 = I \qquad\qquad \{\sigma_i, \sigma_j\} = 2\delta_{ij}$$

Thus, using Eq.(210), Eq.(209) can be written as:

$$\left[(p^0)^2 - |\vec{p}|^2\right] \phi_L(\vec{p}) = m^2 \phi_L(\vec{p}) \tag{211}$$

Eq.(211) suggests that $\phi_L(\vec{p})$ satisfies the Klein-Gordon equation. We can show that $\phi_R(\vec{p})$ also satisfies the Klein-Gordon equation, by eliminating $\phi_L(\vec{p})$ in Eq.(208). Let us go back to Eq.(208):

$$\begin{aligned}(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(\vec{p}) &= m\phi_R(\vec{p})\\ (p^0 - \vec{\sigma}\cdot\vec{p})\phi_R(\vec{p}) &= m\phi_L(\vec{p})\end{aligned}$$

From the energy-momentum relation,

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

we can express p^0 and $|\vec{p}|$ as follows:

$$p^0 = m\cosh\eta > 0 \qquad\qquad |\vec{p}| = m\sinh\eta \tag{212}$$

We are specifying $p^0 > 0$, because the energy of a particle can be negative as we discussed before! Then $\vec{\sigma}\cdot\vec{p}$ can be written as:

$$\vec{\sigma}\cdot\vec{p} = \vec{\sigma}\cdot|\vec{p}|\hat{p} = |\vec{p}|\vec{\sigma}\cdot\hat{p} = m\sinh\eta(\vec{\sigma}\cdot\hat{p}) \tag{213}$$

$(\vec{\sigma}\cdot\hat{p})^2 = 1$ because we have $\hat{p}$, instead of $\vec{p}$ (See the last two steps in Eq.(210)). Then we have the following relations:

$$(\vec{\sigma}\cdot\hat{p})^{\text{even number}} = 1 \qquad\qquad (\vec{\sigma}\cdot\hat{p})^{\text{odd number}} = \vec{\sigma}\cdot\hat{p} \tag{214}$$

Then the operator $(p^0 \pm \vec{\sigma}\cdot\vec{p})$ in Eq.(208) can be written as follows:

$$\begin{aligned}p^0 \pm \vec{\sigma}\cdot\vec{p} &= m\cosh\eta \pm m\sinh\eta(\vec{\sigma}\cdot\hat{p})\\ &= m(\cosh\eta \pm \sinh\eta(\vec{\sigma}\cdot\hat{p}))\\ &= m\left[\left(1 + \frac{1}{2!}\eta^2 + \frac{1}{4!}\eta^4 + \cdots\right) \pm (\vec{\sigma}\cdot\hat{p})\left(\frac{1}{1!}\eta^1 + \frac{1}{3!}\eta^3 + \frac{1}{5!}\eta^5 \cdots\right)\right]\\ &= m\left(1 + \frac{1}{2!}\eta^2(\vec{\sigma}\cdot\hat{p})^2 + \frac{1}{4!}\eta^4(\vec{\sigma}\cdot\hat{p})^4 + \cdots\right)\\ &\pm m\left(\frac{1}{1!}\eta^1(\vec{\sigma}\cdot\hat{p})^1 + \frac{1}{3!}\eta^3(\vec{\sigma}\cdot\hat{p})^3 + \frac{1}{5!}\eta^5(\vec{\sigma}\cdot\hat{p})^5 + \cdots\right)\\ &= m\left(1 \pm \frac{1}{1!}(\vec{\sigma}\cdot\eta\hat{p})^1 + \frac{1}{2!}(\vec{\sigma}\cdot\eta\hat{p})^2 \pm \frac{1}{3!}(\vec{\sigma}\cdot\eta\hat{p})^3 + \frac{1}{4!}(\vec{\sigma}\cdot\eta\hat{p})^4 + \cdots\right)\\ &= me^{\pm\vec{\sigma}\cdot\vec{\eta}}\end{aligned} \tag{215}$$

where $\vec{\eta} = \eta\hat{p}$. Then Eq.(208) can be written as:

$$\begin{aligned} me^{\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) &= m\phi_R(\vec{p}) \\ me^{-\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) &= m\phi_L(\vec{p}) \end{aligned} \tag{216}$$

By canceling out m,

$$\begin{aligned} e^{\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) &= \phi_R(\vec{p}) \\ e^{-\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) &= \phi_L(\vec{p}) \end{aligned} \tag{217}$$

By multiplying $e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}$ to the first equation of Eq.(217) and $e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}$ to the second, both from the left,

$$\begin{aligned} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) &= e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) &= e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) \end{aligned} \tag{218}$$

To solve Eq.(218), we can define $\phi_L(\vec{p})$ and $\phi_R(\vec{p})$ as follows:

$$\phi_L(\vec{p}) = N\sqrt{m}e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \qquad\qquad \phi_R(\vec{p}) = N\sqrt{m}e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \tag{219}$$

So its four-component solution can be written as:

$$u_r(p) = N\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\xi_r\\ \xi_r\end{pmatrix} \tag{220}$$

So $\psi(x)$ can be expressed as (See Eq.(207)):

$$\psi(x) = Nu_r(p)e^{-ip\cdot x} = N^2\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\xi_r\\ \xi_r\end{pmatrix}e^{-ip\cdot x} \tag{221}$$

By Eq.(181), we see that the solution is just a Lorentz-transformed spinor, with $\vec{\theta} = 0$.

From Eq.(212), we assumed $E = p^0 > 0$. Then the natural question is, "how about $E = p^0 < 0$?" Let us take a moment on how we obtained the negative energy. Section 1.7.2 ends with the following sentence:

> The negative energy of a particle corresponds to the particle traveling backward in time, and this is the same as a particle with an opposite charge (a.k.a. antiparticle) traveling forward in time.

We are going to focus on 'The negative energy of a particle corresponds to the particle traveling *backward* in time.' According to this statement, there are three things going on:

$$t \to -t \qquad\qquad E \to -E \qquad\qquad \vec{p} = m\frac{d\vec{x}}{dt} \to m\frac{d\vec{x}}{d(-t)} = -\vec{p} \tag{222}$$

By Eq.(222), Eq.(208) changes into the following:

$$\begin{aligned} &-(p^0 + \vec{\sigma}\cdot\vec{p})\phi_L(\vec{p}) = m\phi_R(\vec{p}) \\ &-(p^0 - \vec{\sigma}\cdot\vec{p})\phi_R(\vec{p}) = m\phi_L(\vec{p}) \end{aligned} \tag{223}$$

By using Eq.(215), Eq.(223) can be written as follows:

$$\begin{aligned} &-me^{\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) = m\phi_R(\vec{p}) \\ &-me^{-\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) = m\phi_L(\vec{p}) \end{aligned} \tag{224}$$

By canceling out m,

$$\begin{aligned} &-e^{\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) = \phi_R(\vec{p}) \\ &-e^{-\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) = \phi_L(\vec{p}) \end{aligned} \tag{225}$$

By multiplying $e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}$ to the first equation of Eq.(225) and $e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}$ to the second, both from the left,

$$\begin{aligned} &-e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) = e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) \\ &-e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_R(\vec{p}) = e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\phi_L(\vec{p}) \end{aligned} \tag{226}$$

To solve Eq.(226), we can define $\phi_L(\vec{p})$ and $\phi_R(\vec{p})$ as follows:

$$\phi_L(\vec{p}) = N\sqrt{m}e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \qquad\qquad \phi_R(\vec{p}) = -N\sqrt{m}e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \tag{227}$$

So its four-component solution can be written as:

$$\begin{aligned} &v_r(p) = N\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\eta_r \\ -\eta_r\end{pmatrix} \\ &\psi(x) = Nv_r(p)e^{ip\cdot x} = N^2\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\eta_r \\ -\eta_r\end{pmatrix}e^{ip\cdot x} \end{aligned} \tag{228}$$

where

$$\eta_r = i\sigma_2\xi_r^* = \begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix}\xi_r^* \tag{229}$$

The plane-wave factor $e^{ip\cdot x}$ comes from applying the time-reversal operation on $e^{-ip\cdot x}$. Notice that $\xi_r^\dagger\eta_r = 0$ because

$$\begin{aligned} \xi_r^\dagger\eta_r = \xi_r^\dagger\begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix}\xi_r^* &= \begin{pmatrix}A^* & B^*\end{pmatrix}\begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix}\begin{pmatrix}A^* \\ B^*\end{pmatrix} \\ &= \begin{pmatrix}A^* & B^*\end{pmatrix}\begin{pmatrix}B^* \\ -A^*\end{pmatrix} = A^*B^* - B^*A^* = 0 \end{aligned} \tag{230}$$

This means the positive energy state and the negative energy state are mutually exclusive. Notice that $u_r^\dagger v_s \neq 0$ because

$$\begin{aligned} u_r^\dagger v_s &= \sqrt{m}\begin{pmatrix}e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger\end{pmatrix}\sqrt{m}\begin{pmatrix}e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s\end{pmatrix} \\ &= m(e^{-\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger\eta_s - e^{\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger\eta_s) \\ &= m(e^{-\vec{\sigma}\cdot\vec{\eta}} - e^{\vec{\sigma}\cdot\vec{\eta}})\xi_r^\dagger\eta_s \neq 0 \end{aligned} \tag{231}$$

because $\xi_r^\dagger \eta_s \neq 0$, generally.

1.14 $u(p)$ and $v(p)$ Gymnastics

Here we are going to take a pause, play around with $u(p)$ and $v(p)$, get familiar with them, and produce some identities because they will be handy in the future.

Identity 1.1

$$\begin{aligned}
\bar{u}_r u_s = u_r^\dagger \gamma^0 u_s &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \end{pmatrix} \\
&= m\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\begin{pmatrix} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \end{pmatrix} \\
&= m(2\xi_r^\dagger \xi_s) = 2m\delta_{rs}
\end{aligned} \tag{232}$$

Identity 1.2

$$\begin{aligned}
u_r^\dagger u_s &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \end{pmatrix} \\
&= m\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_s \end{pmatrix} \\
&= m(e^{-\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \xi_s + e^{\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \xi_s) = (me^{-\vec{\sigma}\cdot\vec{\eta}} + me^{\vec{\sigma}\cdot\vec{\eta}})\xi_r^\dagger \xi_s \\
&= (p^0 - \vec{\sigma}\cdot\vec{p} + p^0 + \vec{\sigma}\cdot\vec{p})\xi_r^\dagger \xi_s \\
&= 2E\delta_{rs}
\end{aligned} \tag{233}$$

where we used $p^0 \pm \vec{\sigma}\cdot\vec{p} = me^{\pm\vec{\sigma}\cdot\vec{\eta}}$ (See Eq.(215)).

Identity 2.1

$$\begin{aligned}
\bar{v}_r v_s = v_r^\dagger \gamma^0 v_s &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger & -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\
&= m\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger & -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \end{pmatrix}\begin{pmatrix} -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\
&= m(-2\eta_r^\dagger \eta_s) = -2m\delta_{rs}
\end{aligned} \tag{234}$$

Identity 2.2

$$\begin{aligned}
v_r^\dagger v_s &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger & -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\
&= m\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger & -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \end{pmatrix}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\
&= m(e^{-\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \eta_s + e^{\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \eta_s) = (me^{-\vec{\sigma}\cdot\vec{\eta}} + me^{\vec{\sigma}\cdot\vec{\eta}})\eta_r^\dagger \eta_s \\
&= (p^0 - \vec{\sigma}\cdot\vec{p} + p^0 + \vec{\sigma}\cdot\vec{p})\eta_r^\dagger \eta_s \\
&= 2E\delta_{rs}
\end{aligned} \tag{235}$$

Identity 3.1

$$\begin{aligned}\bar{u}_r v_s = u_r^\dagger \gamma^0 v_s &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\ &= m\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\begin{pmatrix} -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\ &= m(-\xi_r^\dagger\eta_s + \xi_r^\dagger\eta_s) = 0\end{aligned} \tag{236}$$

Identity 3.2

$$\bar{v}_r u_s = v_r^\dagger \gamma^0 u_s = (u_s^\dagger \gamma^0 v_r)^\dagger = (\bar{u}_s v_r)^\dagger = 0 \tag{237}$$

Identity 4.1

$$\begin{aligned}u_r^\dagger(\vec{p})v_s(-\vec{p}) &= \sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix}\sqrt{m}\begin{pmatrix} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \\ -e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_s \end{pmatrix} \\ &= m(\xi_r^\dagger\eta_s - \xi_r^\dagger\eta_s) = 0\end{aligned} \tag{238}$$

Identity 4.2

$$v_r^\dagger(-\vec{p})u_s(\vec{p}) = [u_s^\dagger(\vec{p})v_r(-\vec{p})]^\dagger = 0 \tag{239}$$

Identity 5.1

$$\begin{aligned}(i\gamma^\mu\partial_\mu - m)\psi(x) &= (i\gamma^\mu\partial_\mu - m)Nu_r(p)e^{-ip\cdot x} \\ &= (i\gamma^0\partial_0 + i\gamma^k\partial_k - m)Nu_r(p)e^{-i(p^0x^0 - p^1x^1 - p^2x^2 - p^3x^3)} \\ &= (\gamma^0 p^0 - \gamma^k p^k - m)Nu_r(p)e^{-ip\cdot x} \\ &= (\gamma^0 p_0 + \gamma^k p_k - m)Nu_r(p)e^{-ip\cdot x} \\ &= (\gamma^\mu p_\mu - m)Nu_r(p)e^{-ip\cdot x} = 0 \\ \\ &\to (\gamma^\mu p_\mu - m)u_r(p) = (\not{p} - m)u_r(p) = 0\end{aligned} \tag{240}$$

where we defined $\gamma^\mu p_\mu = \not{p}$.

Identity 5.2

$$\begin{aligned}[(\gamma^\mu p_\mu - m)u_r(p)]^\dagger &= u_r(p)^\dagger(\gamma^\mu p_\mu - m)^\dagger \\ &= u_r(p)^\dagger(\gamma^0 p_0 + \gamma^k p_k - m)^\dagger \\ &= u_r(p)^\dagger[(\gamma^0)^\dagger p_0 + (\gamma^k)^\dagger p_k - m] \\ &= u_r(p)^\dagger[\gamma^0 p_0 - \gamma^k p_k - m] = 0\end{aligned} \tag{241}$$

where we used $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^k)^\dagger = -\gamma^k$ (See Eq.(112)). By multiplying γ^0 from the right,

$$\begin{aligned}
[(\gamma^\mu p_\mu - m)u_r(p)]^\dagger &= u_r(p)^\dagger[\gamma^0 p_0 - \gamma^k p_k - m]\gamma^0 \\
&= u_r(p)^\dagger[\gamma^0\gamma^0 p_0 - \gamma^k\gamma^0 p_k - m\gamma^0] \\
&= u_r(p)^\dagger[\gamma^0\gamma^0 p_0 + \gamma^0\gamma^k p_k - m\gamma^0] \\
&= u_r(p)^\dagger\gamma^0[\gamma^0 p_0 + \gamma^k p_k - m] \\
&= \bar{u}_r(p)[\gamma^\mu p_\mu - m] = 0 \\
\\
&\rightarrow \bar{u}_r(p)(\gamma^\mu p_\mu - m) = \bar{u}_r(p)(\not{p} - m) = 0
\end{aligned} \tag{242}$$

where we used $\{\gamma^k, \gamma^0\} = 0$ (Eq.(107)).

Identity 6.1

$$\begin{aligned}
(i\gamma^\mu\partial_\mu - m)\psi(x) &= (i\gamma^\mu\partial_\mu - m)Nv_r(p)e^{ip\cdot x} \\
&= (i\gamma^0\partial_0 + i\gamma^k\partial_k - m)Nv_r(p)e^{i(p^0x^0 - p^1x^1 - p^2x^2 - p^3x^3)} \\
&= (-\gamma^0 p^0 + \gamma^k p^k - m)Nv_r(p)e^{ip\cdot x} \\
&= (-\gamma^0 p_0 - \gamma^k p_k - m)Nv_r(p)e^{ip\cdot x} \\
&= (-\gamma^\mu p_\mu - m)Nv_r(p)e^{ip\cdot x} = 0 \\
\\
&\rightarrow (\gamma^\mu p_\mu + m)v_r(p) = (\not{p} + m)v_r(p) = 0
\end{aligned} \tag{243}$$

Notice that we used $e^{ip\cdot x}$, rather than $e^{-ip\cdot x}$. It is because $p^0 \rightarrow -p^0$ and $p^k \rightarrow -p^k$ that we time reverse to deal with negative energy, as discussed in Section 1.7.2.

Identity 6.2

$$\begin{aligned}
[(\gamma^\mu p_\mu + m)v_r(p)]^\dagger &= v_r(p)^\dagger(\gamma^\mu p_\mu + m)^\dagger \\
&= v_r(p)^\dagger(\gamma^0 p_0 + \gamma^k p_k + m)^\dagger \\
&= v_r(p)^\dagger[(\gamma^0)^\dagger p_0 + (\gamma^k)^\dagger p_k + m] \\
&= v_r(p)^\dagger[\gamma^0 p_0 - \gamma^k p_k + m] = 0
\end{aligned} \tag{244}$$

where we used $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^k)^\dagger = -\gamma^k$ (See Eq.(111)). By multiplying γ^0 from the right,

$$\begin{aligned}
[(\gamma^\mu p_\mu + m)v_r(p)]^\dagger &= v_r(p)^\dagger[\gamma^0 p_0 - \gamma^k p_k + m]\gamma^0 \\
&= v_r(p)^\dagger[\gamma^0\gamma^0 p_0 - \gamma^k\gamma^0 p_k + m\gamma^0] \\
&= v_r(p)^\dagger[\gamma^0\gamma^0 p_0 + \gamma^0\gamma^k p_k + m\gamma^0] \\
&= v_r(p)^\dagger\gamma^0[\gamma^0 p_0 + \gamma^k p_k + m] \\
&= \bar{v}_r(p)[\gamma^\mu p_\mu + m] = 0 \\
\\
&\to \bar{v}_r(p)(\gamma^\mu p_\mu + m) = \bar{v}_r(p)(\not{p} + m) = 0
\end{aligned} \tag{245}$$

where we used $\{\gamma^k, \gamma^0\} = 0$ (Eq.(107)).

Identities 5.1, 5.2, 6.1, and 6.2 involve $(\not{p} + m)$ or $(\not{p} - m)$. Let's investigate the meaning of the operators.

Identity 7.1

$$\begin{aligned}
\Lambda_+ = \frac{1}{2m}\sum_{r=1,2} u_r(p)\bar{u}_r(p) &= \frac{1}{2m}\sum_{r=1,2} N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \end{pmatrix} N\sqrt{m}\begin{pmatrix} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger & e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r^\dagger \end{pmatrix} \\
&= \frac{1}{2m}N^2 m\sum_{r=1,2}\begin{pmatrix} \xi_r\xi_r^\dagger & e^{-\vec{\sigma}\cdot\vec{\eta}}\xi_r\xi_r^\dagger \\ e^{\vec{\sigma}\cdot\vec{\eta}}\xi_r\xi_r^\dagger & \xi_r\xi_r^\dagger \end{pmatrix}
\end{aligned} \tag{246}$$

Notice that

$$\xi_1\xi_1^\dagger = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \xi_2\xi_2^\dagger = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Thus,

$$\xi_1\xi_1^\dagger + \xi_2\xi_2^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \tag{247}$$

With Eq.(247), we can proceed calculating Eq.(246):

$$\begin{aligned}
\Lambda_+ &= \frac{1}{2m}N^2 m \sum_{r=1,2} \begin{pmatrix} \xi_r\xi_r^\dagger & e^{-\vec{\sigma}\cdot\vec{\eta}}\xi_r\xi_r^\dagger \\ e^{\vec{\sigma}\cdot\vec{\eta}}\xi_r\xi_r^\dagger & \xi_r\xi_r^\dagger \end{pmatrix} = \frac{1}{2m}N^2 m \begin{pmatrix} 1 & e^{-\vec{\sigma}\cdot\vec{\eta}} \\ e^{\vec{\sigma}\cdot\vec{\eta}} & 1 \end{pmatrix} \\
&= \frac{1}{2m}N^2 \begin{pmatrix} m & p^0 - \vec{\sigma}\cdot\vec{p} \\ p^0 + \vec{\sigma}\cdot\vec{p} & m \end{pmatrix} \\
&= \frac{1}{2m}\left[\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} + \begin{pmatrix} 0 & p^0 \\ p^0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -\sigma_k p^k \\ \sigma_k p^k & 0 \end{pmatrix}\right] \\
&= \frac{1}{2m}\left[\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p^0 + \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}(-p^k)\right] \\
&= \frac{1}{2m}\left[\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_0 + \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}(p_k)\right] \\
&= \frac{1}{2m}(m + \gamma^0 p_0 + \gamma^k p_k) \\
&= \frac{m + \gamma^\mu p_\mu}{2m} = \frac{m + \not{p}}{2m}
\end{aligned} \tag{248}$$

Let us start with Eq.(240):

$$(\not{p} - m)u_r(p) = 0$$

We can add $2mu_r(p)$ on both sides:

$$(\not{p} + m)u_r(p) = 2mu_r(p) \tag{249}$$

By dividing both sides by $\frac{1}{2m}$,

$$\frac{\not{p} + m}{2m}u_r(p) = u_r(p) \tag{250}$$

Thus, we have:

$$\frac{(\not{p} - m)}{2m}u_r(p) = 0 \qquad\qquad \frac{\not{p} + m}{2m}u_r(p) = u_r(p) \tag{251}$$

Hmm... What does this mean? To answer this question, we define another operator Λ_-. To do so, keep in mind Eq.(229):

$$\eta_r = i\sigma_2\xi_r^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\xi_r^*$$

Identity 7.2

$$\begin{aligned}
\Lambda_- = -\frac{1}{2m}\sum_{r=1,2} v_r(p)\bar{v}_r(p) &= -\frac{1}{2m}\sum_{r=1,2} N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix} N\sqrt{m}\begin{pmatrix} -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger & e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r^\dagger \end{pmatrix} \\
&= -\frac{1}{2m}N^2 m \sum_{r=1,2}\begin{pmatrix} -\eta_r\eta_r^\dagger & e^{-\vec{\sigma}\cdot\vec{\eta}}\eta_r\eta_r^\dagger \\ e^{\vec{\sigma}\cdot\vec{\eta}}\eta_r\eta_r^\dagger & -\eta_r\eta_r^\dagger \end{pmatrix}
\end{aligned} \tag{252}$$

Notice that

$$\eta_1\eta_1^\dagger = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \eta_2\eta_2^\dagger = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus,

$$\eta_1\eta_1^\dagger + \eta_2\eta_2^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \tag{253}$$

With Eq.(247), we can proceed calculating Eq.(246):

$$\begin{aligned}
\Lambda_- &= -\frac{1}{2m}N^2 m \sum_{r=1,2} \begin{pmatrix} -\eta_r\eta_r^\dagger & e^{-\vec{\sigma}\cdot\vec{\eta}}\eta_r\eta_r^\dagger \\ e^{\vec{\sigma}\cdot\vec{\eta}}\eta_r\eta_r^\dagger & -\eta_r\eta_r^\dagger \end{pmatrix} = -\frac{1}{2m}N^2 m \begin{pmatrix} -1 & e^{-\vec{\sigma}\cdot\vec{\eta}} \\ e^{\vec{\sigma}\cdot\vec{\eta}} & -1 \end{pmatrix} \\
&= -\frac{1}{2m}N^2 \begin{pmatrix} -m & p^0 - \vec{\sigma}\cdot\vec{p} \\ p^0 + \vec{\sigma}\cdot\vec{p} & -m \end{pmatrix} \\
&= -\frac{1}{2m}\left[\begin{pmatrix} -m & 0 \\ 0 & -m \end{pmatrix} + \begin{pmatrix} 0 & p^0 \\ p^0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -\sigma_k p^k \\ \sigma_k p^k & 0 \end{pmatrix} \right] \\
&= -\frac{1}{2m}\left[\begin{pmatrix} -m & 0 \\ 0 & -m \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p^0 + \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} (-p^k) \right] \\
&= -\frac{1}{2m}\left[\begin{pmatrix} -m & 0 \\ 0 & -m \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_0 + \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} (p_k) \right] \\
&= -\frac{1}{2m}(-m + \gamma^0 p_0 + \gamma^k p_k) \\
&= -\frac{-m + \gamma^\mu p_\mu}{2m} = -\frac{-m + \not{p}}{2m}
\end{aligned} \tag{254}$$

Let us start with Eq.(243):

$$(\not{p} + m)v_r(p) = 0$$

We can subtract $2mv_r(p)$ on both sides:

$$(\not{p} - m)v_r(p) = -2mv_r(p) \tag{255}$$

By dividing both sides by $-\frac{1}{2m}$,

$$\frac{\not{p} - m}{-2m}v_r(p) = v_r(p) \tag{256}$$

Thus, we have:

$$\frac{(\not{p} + m)}{2m}v_r(p) = 0 \qquad \frac{\not{p} - m}{-2m}v_r(p) = v_r(p) \tag{257}$$

Rewriting Eq.(251) and Eq.(257),

$$\frac{(\not{p} - m)}{-2m}u_r(p) = \Lambda_- u_r(p) = 0 \qquad \frac{\not{p} + m}{2m}u_r(p) = \Lambda_+ u_r(p) = u_r(p)$$

$$\frac{(\not{p}+m)}{2m}v_r(p) = \Lambda_+ v_r(p) = 0 \qquad \frac{\not{p}-m}{-2m}v_r(p) = \Lambda_- v_r(p) = v_r(p)$$

Thus, Λ_+ (Λ_-) is an operator which projects a spinor onto $u_r(p)$ ($v_r(p)$). In other words, Λ_+ (Λ_-) picks the positive (negative) energy solution. Notice that Λ_- can be obtained from Λ_+ by taking $m \to -m$, which makes $E = mc^2 \to E = -mc^2$.

1.15 Canonical Momentum

We would like to calculate the magnetic moment of an electron with the Dirac equation. However, before doing so, I would like to introduce canonical momentum, which arises in electrodynamics. Let us start with Maxwell's equations:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad \nabla \cdot \vec{B} = 0$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}$$

From $\nabla \cdot \vec{B} = 0$, we can let $\vec{B} = \nabla \times \vec{A}$, because the divergence of the curl of a vector field is always zero. Then the curl of the electric field can be written as:

$$\begin{aligned} \nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} = -\frac{\partial}{\partial t}(\nabla \times \vec{A}) \\ &\to \nabla \times (\vec{E} + \frac{\partial \vec{A}}{\partial t}) = 0 \end{aligned} \tag{258}$$

By using the fact that the curl of the gradient of a vector is always zero,

$$-\nabla V = \vec{E} + \frac{\partial \vec{A}}{\partial t} \tag{259}$$

Thus the electric and magnetic fields can be written as:

$$\vec{E} = -\nabla V - \frac{\partial \vec{A}}{\partial t} \qquad \vec{B} = \nabla \times \vec{A} \tag{260}$$

Let us continue with expressing the Lorentz force law in terms of the electromagnetic potentials:

$$\vec{F} = \frac{d\vec{p}}{dt} = q(\vec{E} + \vec{v} \times \vec{B}) = q\left[-\nabla V - \frac{\partial \vec{A}}{\partial t} + \vec{v} \times (\nabla \times \vec{A})\right] \tag{261}$$

By using the following vector identity,

$$\nabla(\vec{v} \cdot \vec{A}) = \vec{v} \times (\nabla \times \vec{A}) + (\vec{v} \cdot \nabla)\vec{A} \tag{262}$$

Eq.(261) can be expressed as:

$$\frac{d\vec{p}}{dt} = q\left[-\nabla V - \frac{\partial \vec{A}}{\partial t} + \nabla(\vec{v}\cdot\vec{A}) - (\vec{v}\cdot\nabla)\vec{A}\right] \tag{263}$$

Notice that $\frac{\partial \vec{A}}{\partial t} + (\vec{v}\cdot\nabla)\vec{A}$ can be simplified by using the chain rule.

$$\begin{aligned}\frac{\partial \vec{A}}{\partial t} + (\vec{v}\cdot\nabla)\vec{A} &= \frac{\partial \vec{A}}{\partial t} + \left(\frac{dx}{dt}\frac{\partial}{\partial x} + \frac{dy}{dt}\frac{\partial}{\partial y} + \frac{dz}{dt}\frac{\partial}{\partial z}\right)\vec{A} \\ &= \frac{d\vec{A}}{dt}\end{aligned} \tag{264}$$

Let us continue with Eq.(263).

$$\frac{d\vec{p}}{dt} = q\left[-\nabla V + \nabla\left(\vec{v}\cdot\vec{A}\right) - \frac{d\vec{A}}{dt}\right] \tag{265}$$

Thus,

$$\frac{d}{dt}\left(\vec{p} + q\vec{A}\right) = -\nabla\left[q\left(V - \vec{v}\cdot\vec{A}\right)\right] \tag{266}$$

Thus we can define a new kind of momentum, called canonical momentum, whose time derivative is given by the negative gradient of $q\left(V - \vec{v}\cdot\vec{A}\right)$:

$$\vec{p}_{\text{canonical}} = \vec{p} + q\vec{A} \tag{267}$$

Hence, to take an electromagnetic interaction between a charged particle and the external electromagnetic fields, we replace the particle's mechanical momentum with the canonical momentum.

1.16 Magnetic Moment of the Electron

Now let us start with the Dirac equation:

$$\begin{aligned}(i\gamma^\mu\partial_\mu - m)\psi(x) &= (i\gamma^0\partial_0 + i\gamma^k\partial_k - m)\psi(x) \\ &= (\gamma^0(i\partial_0) - \gamma^k(-i\partial_k) - m)\psi(x) \\ &= (\gamma^0 p_0 - \gamma^k p_k - m)\psi(x)\end{aligned}$$

With Eq.(267), $q = -e$, $p_\mu = (p_0, p_1, p_2, p_3)$, and $A_\mu = (A_0, -\vec{A})$,

$$\begin{aligned}&\xrightarrow{p_\mu = p_\mu + eA_\mu} (\gamma^0(p_0 + eA_0) - \gamma^k(p_k + eA_k) - m)\psi(x) \\ &= \left[(i\partial_0 + eA_0)\gamma^0 - (-i\partial_k + eA_k)\gamma^k - m\right]\psi(x) \\ &= \begin{pmatrix} -m & i\partial_0 + eA_0 - \vec{\sigma}\cdot(-i\nabla + e\vec{A}) \\ i\partial_0 + eA_0 + \vec{\sigma}\cdot(-i\nabla + e\vec{A}) & -m \end{pmatrix}\begin{pmatrix}\psi_1 \\ \psi_2\end{pmatrix} = 0\end{aligned}$$

where we used

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}$$

Thus, the expanded equation is:

$$\begin{aligned} \left[i\partial_0 + eA_0 - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\psi_2 &= m\psi_1 \\ \left[i\partial_0 + eA_0 + \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\psi_1 &= m\psi_2 \end{aligned} \tag{268}$$

With

$$A_0 \to \Phi \tag{269}$$

Eq.(268) can be rewritten as:

$$\begin{aligned} \left[i\partial_0 + e\Phi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\psi_2 &= m\psi_1 \\ \left[i\partial_0 + e\Phi + \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\psi_1 &= m\psi_2 \end{aligned} \tag{270}$$

For the high-energy limit ($E \approx p$), we can ignore the mass term, which decouples the two equations in Eq.(270). Then Eq.(270) becomes the Weyl equations under the electromagnetic fields (the mechanical momentum is replaced with the canonical momentum).

However, consider the non-relativistic limit, which is $m >> |\vec{p}|, \Phi, ...$. Then $i\partial_0 \sim E \approx m$. Then Eq.(270) becomes:

$$\begin{aligned} m\psi_2 &\approx m\psi_1 \\ m\psi_1 &\approx m\psi_2 \end{aligned} \tag{271}$$

Thus, in the non-relativistic limit,

$$\psi_1 \approx \psi_2 \tag{272}$$

So it would be helpful to introduce

$$\psi_+ = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2) \qquad\qquad \psi_- = \frac{1}{\sqrt{2}}(-\psi_1 + \psi_2) \tag{273}$$

In the non-relativistic limit, $\psi_- \approx 0$. Eq.(273) can be solved for ψ_1 and ψ_2.

$$\psi_1 = \frac{1}{\sqrt{2}}(\psi_+ - \psi_-) \qquad\qquad \psi_2 = \frac{1}{\sqrt{2}}(\psi_+ + \psi_-) \tag{274}$$

By plugging Eq.(274) into Eq.(270),

$$\begin{aligned} \left[i\partial_0 + e\Phi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\frac{1}{\sqrt{2}}(\psi_+ + \psi_-) &= m\frac{1}{\sqrt{2}}(\psi_+ - \psi_-) \\ \left[i\partial_0 + e\Phi + \vec{\sigma}\cdot(-i\nabla + e\vec{A})\right]\frac{1}{\sqrt{2}}(\psi_+ - \psi_-) &= m\frac{1}{\sqrt{2}}(\psi_+ + \psi_-) \end{aligned} \tag{275}$$

By expanding Eq.(275),

$$(i\partial_0 + e\Phi)\psi_+ - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_+ + (i\partial_0 + e\Phi)\psi_- - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_- = m(\psi_+ - \psi_-) \qquad \text{(i)}$$

$$(i\partial_0 + e\Phi)\psi_+ + \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_+ - (i\partial_0 + e\Phi)\psi_- - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_- = m(\psi_+ + \psi_-) \qquad \text{(ii)}$$

(i) + (ii)

$$(i\partial_0 + e\Phi)\psi_+ - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_- = m\psi_+ \tag{276}$$

(i) - (ii)

$$(i\partial_0 + e\Phi)\psi_- - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\psi_+ = -m\psi_- \tag{277}$$

Because $E = m + T$ where T is kinetic energy, the time-dependent factor of the solution of the Dirac equation is $e^{-i(m+T)t}$. Thus we can write ψ_+ and ψ_- as:

$$\psi_+ = \phi e^{-i(m+T)t} \qquad\qquad \psi_- = \chi e^{-i(m+T)t} \tag{278}$$

By plugging Eq.(278) into Eq.(276) and Eq.(277),

$$\begin{aligned} (m + T + e\Phi)\phi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\chi &= m\phi \\ (m + T + e\Phi)\chi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\phi &= -m\chi \end{aligned} \tag{279}$$

By using the non-relativistic limit on the second equation above (T, $e\Phi << m$),

$$\begin{aligned} &(m + T + e\Phi)\phi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\chi = m\phi \qquad \text{(i)} \\ &m\chi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\phi = -m\chi \qquad \text{(ii)} \end{aligned} \tag{280}$$

Solving for χ,

$$\chi = \frac{1}{2m}\vec{\sigma}\cdot(-i\nabla + e\vec{A})\phi \tag{281}$$

By plugging Eq.(281) into Eq.(280(i)),

$$(m + T + e\Phi)\phi - \vec{\sigma}\cdot(-i\nabla + e\vec{A})\frac{1}{2m}\vec{\sigma}\cdot(-i\nabla + e\vec{A})\phi = m\phi \tag{282}$$

By solving for $T\phi$,

$$T\phi = \left(-e\Phi + \frac{1}{2m}\vec{\sigma}\cdot(-i\nabla + e\vec{A})\vec{\sigma}\cdot(-i\nabla + e\vec{A})\right)\phi \tag{283}$$

Using the following identity,

$$
\begin{aligned}
\left(\vec{\sigma}\cdot\vec{C}\right)\left(\vec{\sigma}\cdot\vec{D}\right) &= \left(\sigma_1 C_1+\sigma_2 C_2+\sigma_3 C_3\right)\left(\sigma_1 D_1+\sigma_2 D_2+\sigma_3 D_3\right)\\
&= \sigma_1^2 C_1 D_1+\sigma_1\sigma_2 C_1 D_2+\sigma_1\sigma_3 C_1 D_3\\
&+ \sigma_2\sigma_1 C_2 D_1+\sigma_2^2 C_2 D_2+\sigma_2\sigma_3 C_2 D_3\\
&+ \sigma_3\sigma_1 C_3 D_1+\sigma_3\sigma_2 C_3 D_2+\sigma_3^2 C_3 D_3\\
&= C_1 D_1+\sigma_1\sigma_2 C_1 D_2-\sigma_3\sigma_1 C_1 D_3\\
&- \sigma_1\sigma_2 C_2 D_1+C_2 D_2+\sigma_2\sigma_3 C_2 D_3\\
&+ \sigma_3\sigma_1 C_3 D_1-\sigma_2\sigma_3 C_3 D_2+C_3 D_3\\
&= \left(C_1 D_1+C_2 D_2+C_3 D_3\right)\\
&+ \sigma_2\sigma_3(C_2 D_3-C_3 D_2)+\sigma_3\sigma_1(C_3 D_1-C_1 D_3)+\sigma_1\sigma_2(C_1 D_2-C_2 D_1)\\
&= \vec{C}\cdot\vec{D}\\
&+ i\sigma_1(C_2 D_3-C_3 D_2)+i\sigma_2(C_3 D_1-C_1 D_3)+i\sigma_3(C_1 D_2-C_2 D_1)\\
&= \vec{C}\cdot\vec{D}+i\vec{\sigma}\cdot\left(\vec{C}\times\vec{D}\right)
\end{aligned}
\tag{284}
$$

Eq.(283) becomes:

$$
T\phi=\left[-e\Phi+\frac{1}{2m}\left(-i\nabla+e\vec{A}\right)\cdot\left(-i\nabla+e\vec{A}\right)+\frac{1}{2m}i\vec{\sigma}\cdot\left(-i\nabla+e\vec{A}\right)\times\left(-i\nabla+e\vec{A}\right)\right]\phi
\tag{285}
$$

Using just one more identity

$$
\begin{aligned}
\left(-i\nabla+e\vec{A}\right)\times\left(-i\nabla+e\vec{A}\right)\phi &= -ie\left(\nabla\times\vec{A}+\vec{A}\times\nabla\right)\phi\\
&= -ie\left(\nabla\times\left(\vec{A}\phi\right)+\vec{A}\times\nabla\phi\right)\\
&= -ie\left(\left(\nabla\times\vec{A}\right)\phi+(\nabla\phi)\times\vec{A}+\vec{A}\times(\nabla\phi)\right)\\
&= -ie\vec{B}\phi
\end{aligned}
\tag{286}
$$

where we used

$$
\nabla\times\nabla\phi=0 \qquad \vec{A}\times\vec{A}=0 \qquad \nabla\times(\vec{F}g)=\left(\nabla\times\vec{F}\right)g+(\nabla g)\times\vec{F}
$$

Finally Eq.(285) can be written as:

$$
T\phi=\left[-e\Phi+\frac{1}{2m}\left(-i\nabla+e\vec{A}\right)^2+\frac{e}{2m}\vec{\sigma}\cdot\vec{B}\right]\phi
\tag{287}
$$

$-e\Phi$ is an electric potential, $\frac{1}{2m}\left(-i\nabla+e\vec{A}\right)^2$ corresponds to the kinetic energy. The last term corresponds to the magnetic interaction term $-\vec{\mu}\cdot\vec{B}$. Thus,

$$
\vec{\mu}=-\frac{e}{2m}\vec{\sigma}=2\cdot\frac{-e}{2m}\vec{s}=g\cdot\frac{-e}{2m}\vec{s}
\tag{288}
$$

The factor 2 is called $g = 2$.

1.16.1 Why is $g = 2$ special?

To answer this question, let us take a look at what the classical theory predicts. Electrodynamics defines a magnetic moment μ as:

$$\mu = IA \tag{289}$$

where I is current and A is an area that the current encloses. Suppose an electron is in a circular orbit of radius r with its speed v and period T. Then the magnetic moment is calculated by:

$$\begin{aligned}\mu = IA &= \left(\frac{q}{T}\right) \cdot (\pi r^2) = \left(q \frac{v}{2\pi r}\right) \cdot (\pi r^2) \\ &= \frac{qvr}{2} = \frac{qmvr}{2m} = \frac{qL}{2m}\end{aligned} \tag{290}$$

where $L = mvr$, which is the angular momentum.

Comparing Eq.(288) and Eq.(290), we see that the classical theory predicts $g = 1$.

1.17 Parity

Although Eq.(130) is discussed in the context of rotation, the lesson is: if you change your input ($x \to x'$), the output also changes ($\psi(x) \to \psi'(x')$). In this section, we want to change our input, x^μ, in the following way:

$$x^\mu = (x^0, \vec{x}) \to x^{\mu'} = (x^0, -\vec{x}) = x_\mu \tag{291}$$

That is, we invert the x, y, and z axes. This transformation is called a parity transformation. The solution of the Dirac equation, $\psi(x)$, is transformed accordingly:

$$\psi(x) \to \psi'(x') = S\psi(x) \tag{292}$$

Then $\bar{\psi}'(x')$ is:

$$\bar{\psi}'(x') = (\psi'(x'))^\dagger \gamma^0 = (S\psi(x))^\dagger \gamma^0 = \psi^\dagger(x) S^\dagger \gamma^0 \tag{293}$$

The parity transformed Dirac equation is:

$$(\gamma^\mu i\partial'_\mu - m)\psi'(x') = 0 \tag{294}$$

Notice, under parity transformation,

$$\partial'_0 = \partial_0 \qquad \partial'_k = -\partial_k \tag{295}$$

Then Eq.(294) can be written as:

$$\begin{aligned}(\gamma^\mu i\partial'_\mu - m)\psi'(x') &= (\gamma^0 i\partial'_0 + \gamma^k i\partial'_k - m)\psi'(x') \\ &= (\gamma^0 i\partial_0 - \gamma^k i\partial_k - m)S\psi(x) = 0\end{aligned} \tag{296}$$

We can insert $SS^{-1} = I$ wherever we want. By multiplying S^{-1} on both sides from the left,

$$\begin{aligned}(\gamma^0 i\partial_0 - \gamma^k i\partial_k - m)S\psi(x) &= 0 \\ S^{-1}(\gamma^0 i\partial_0 - \gamma^k i\partial_k - m)SS^{-1}S\psi(x) &= 0 \\ (S^{-1}\gamma^0 S i\partial_0 - S^{-1}\gamma^k S i\partial_k - m)\psi(x) &= 0\end{aligned} \tag{297}$$

For Eq.(297) to satisfy the Dirac equation, we impose:

$$S^{-1}\gamma^0 S = \gamma^0 \qquad\qquad S^{-1}\gamma^k S = -\gamma^k \tag{298}$$

Then Eq.(297) becomes:

$$(\gamma^0 i\partial_0 + \gamma^k i\partial_k - m)\psi(x) = (\gamma^\mu i\partial_\mu - m)\psi(x) = 0 \tag{299}$$

as we demanded.

We already know what S is, based on Eq.(298). Take a look at Eq.(108(iv)):

$$\begin{aligned}\{\gamma^k, \gamma^0\} = 0 &\rightarrow \gamma^k\gamma^0 + \gamma^0\gamma^k = 0 \\ &\rightarrow \gamma^k\gamma^0 = -\gamma^0\gamma^k \\ &\rightarrow \gamma^k = -\gamma^0\gamma^k\gamma^0\end{aligned} \tag{300}$$

because $(\gamma^0)^{-1} = \gamma^0$.

By comparing Eq.(298) and Eq.(300), we see that

$$S = \gamma^0 \tag{301}$$

Let us apply a parity operation on a particle and an antiparticle.

1. Particle

When we deal with a particle (which is $E > 0$), its state is given by Eq.(220):

$$u_r(\vec{p}) = N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \end{pmatrix}$$

Let's apply a parity operation on $u_r(\vec{p})$.

$$Su_r(\vec{p}) = \gamma^0 u_r(\vec{p}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \end{pmatrix} = \begin{pmatrix} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \end{pmatrix} = u_r(-\vec{p}) \tag{302}$$

The last equality comes from $\vec{\eta} = \eta\hat{p}$. If we apply a parity operation on $u_r(\vec{p})$, it has the same effect as replacing $\vec{\eta}$ with $-\vec{\eta}$. Then, from $\vec{\eta} = \eta\hat{p}$, here the parity operation does $\hat{p} \to -\hat{p}$.

2. Antiparticle

When we deal with an antiparticle (which is $E < 0$), its state is given by Eq.(228):

$$v_r(\vec{p}) = N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix}$$

Let's apply a parity operation on $u_r(\vec{p})$.

$$\begin{aligned} Sv_r(\vec{p}) = \gamma^0 v_r(\vec{p}) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix} \\ &= \begin{pmatrix} -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix} = -\begin{pmatrix} e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ -e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix} = -v_r(-\vec{p}) \end{aligned} \tag{303}$$

Again, the last equality comes from $\vec{\eta} = \eta\hat{p}$.

1.18 Parity Gymnastics

Eq.(298) means something – it seems that under parity γ^μ transforms the same way as a 4-vector.

$$S^{-1}\gamma^0 S = \gamma^0 \qquad\qquad S^{-1}\gamma^k S = -\gamma^k$$

Also, because we know $S = \gamma^0$, we can finalize proceeding Eq.(293):

$$\begin{aligned} \bar{\psi}(x) \to \bar{\psi}'(x') = (\psi'(x'))^\dagger\gamma^0 = (S\psi(x))^\dagger\gamma^0 = \psi^\dagger(x)S^\dagger\gamma^0 \\ = \psi^\dagger(x)\gamma^0 S^{-1} = \bar{\psi}(x)S^{-1} \end{aligned} \tag{304}$$

where we used $\gamma^0 = S = S^\dagger = S^{-1}$ so they can be freely interchanged. So we discovered how $\psi(x)$ or γ^μ transforms under parity:

$$\psi(x) \to \psi'(x') = S\psi(x) \qquad \bar{\psi}(x) \to \bar{\psi}'(x') = \bar{\psi}(x)S^{-1} \qquad \textit{Set 1}$$

$$S^{-1}\gamma^0 S = \gamma^0 \qquad S^{-1}\gamma^k S = -\gamma^k \qquad \textit{Set 2}$$

The "rule" of this specific game is to use *either* Set 1 or Set 2 when we apply a parity transformation. We cannot use *both.* I would like to think of Set 1 as the "Schrodinger picture" where the state changes and Set 2 as the "Heisenberg picture" where the operator changes. Let us apply this rule to some quantities as exercises.

1. Parity of $\bar{\psi}(x)\psi(x)$

$$\bar{\psi}(x)\psi(x) \xrightarrow{\text{parity}} \bar{\psi}'(x')\psi'(x') = \bar{\psi}(x)S^{-1}S\psi(x) = \bar{\psi}\psi(x) \qquad \text{Set 1}$$

Thus,

$$\bar{\psi}(x)\psi(x) \xrightarrow{\text{parity}} \bar{\psi}\psi(x) \tag{305}$$

2. Parity of $i\bar{\psi}(x)\gamma^5\psi(x)$ where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$

$$\begin{aligned}
i\bar{\psi}(x)\gamma^5\psi(x) \xrightarrow{\text{parity}} i\bar{\psi}'(x')\gamma^5\psi'(x') &= i\bar{\psi}(x)S^{-1}i\gamma^0\gamma^1\gamma^2\gamma^3S\psi(x) \qquad \text{Set 1} \\
&= i\bar{\psi}iS^{-1}\gamma^0(SS^{-1})\gamma^1(SS^{-1})\gamma^2(SS^{-1})\gamma^3S\psi(x) \\
&= i\bar{\psi}i(S^{-1}\gamma^0S)(S^{-1}\gamma^1S)(S^{-1}\gamma^2S)(S^{-1}\gamma^3S)\psi(x) \\
&= i\bar{\psi}i(\gamma^0)(-\gamma^1)(-\gamma^2)(-\gamma^3)\psi(x) \qquad \text{Set 2} \\
&= -i\bar{\psi}(x)\gamma^5\psi(x)
\end{aligned}$$

Thus,

$$i\bar{\psi}(x)\gamma^5\psi(x) \xrightarrow{\text{parity}} -i\bar{\psi}(x)\gamma^5\psi(x) \tag{306}$$

From Eq.(306), we can say

$$\gamma^5 \xrightarrow{\text{parity}} -\gamma^5 \qquad \text{or equivalently} \qquad S^{-1}\gamma^5S = -\gamma^5 \tag{307}$$

3. Parity of $\bar{\psi}(x)\gamma^\mu\psi(x)$

$$\begin{aligned}
\bar{\psi}(x)\gamma^\mu\psi(x) \xrightarrow{\text{parity}} \bar{\psi}'(x')\gamma^\mu\psi'(x') &= \bar{\psi}(x)S^{-1}\gamma^\mu S\psi(x) \qquad \text{Set 1} \\
&= \bar{\psi}(x)\gamma_\mu\psi(x) \qquad \text{Set 2}
\end{aligned} \tag{308}$$

where $\gamma^\mu = (\gamma^0, \vec{\gamma})$ and $\gamma_\mu = (\gamma^0, -\vec{\gamma})$ (Forgive my abusing the notation – γ^μ is **not** a 4-vector. However, I am raising/lowering the indices *as if* it is a 4-vector for convenience.) Thus,

$$\bar{\psi}(x')\gamma^\mu\psi(x) \xrightarrow{\text{parity}} \bar{\psi}(x)\gamma_\mu\psi(x) \tag{309}$$

This makes sense. Eq.(198) states that $\bar{\psi}\gamma^\mu\psi$ is a 4-vector. Also, Eq.(55), which is the same as $\bar{\psi}\gamma^\mu\psi$, corresponds to a 4-current (See Eq.(68)). Thus, its spatial components must flip their sign under parity transformation. From Eq.(308), we can say

$$\gamma^\mu \xrightarrow{\text{parity}} \gamma_\mu \qquad \text{or equivalently} \qquad S^{-1}\gamma^\mu S = \gamma_\mu \tag{310}$$

4. Parity of $\bar{\psi}(x)\gamma^5\gamma^\mu\psi(x)$

$$\begin{aligned}
\bar{\psi}(x)\gamma^5\gamma^\mu\psi(x) \xrightarrow{\text{parity}} \bar{\psi}'(x')\gamma^5\gamma^\mu\psi'(x') &= \bar{\psi}(x)S^{-1}\gamma^5\gamma^\mu S\psi(x) \qquad \text{Set 1} \\
&= \bar{\psi}(x)S^{-1}\gamma^5(SS^{-1})\gamma^\mu S\psi(x) \\
&= \bar{\psi}(x)(S^{-1}\gamma^5S)(S^{-1}\gamma^\mu S)\psi(x) \\
&= \bar{\psi}(x)(-\gamma^5)(\gamma_\mu)\psi(x) \\
&= -\bar{\psi}(x)\gamma^5\gamma_\mu\psi(x)
\end{aligned} \tag{311}$$

Thus,

$$\bar{\psi}(x)\gamma^5\gamma^\mu\psi(x) \xrightarrow{\text{parity}} -\bar{\psi}(x)\gamma^5\gamma_\mu\psi(x) \tag{312}$$

From Eq.(312), we can say

$$\gamma^5\gamma^\mu \xrightarrow{\text{parity}} -\gamma^5\gamma_\mu \qquad \text{or equivalently} \qquad S^{-1}\gamma^5\gamma^\mu S = -\gamma^5\gamma_\mu \tag{313}$$

5. Parity of $\bar{\psi}(x)\sigma^{\mu\nu}\psi(x)$ where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$

$$\begin{aligned}
\bar{\psi}(x)\sigma^{\mu\nu}\psi(x) \xrightarrow{\text{parity}} \bar{\psi}'(x')\sigma^{\mu\nu}\psi'(x') &= \bar{\psi}(x)S^{-1}\sigma^{\mu\nu}S\psi(x) \qquad \text{Set 1} \\
&= \bar{\psi}(x)S^{-1}\frac{i}{2}[\gamma^\mu, \gamma^\nu]S\psi(x) \\
&= \bar{\psi}(x)\frac{i}{2}S^{-1}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)S\psi(x) \\
&= \bar{\psi}(x)\frac{i}{2}(S^{-1}\gamma^\mu SS^{-1}\gamma^\nu S - S^{-1}\gamma^\nu SS^{-1}\gamma^\mu S)\psi(x) \qquad \text{Set 2} \\
&= \bar{\psi}(x)\frac{i}{2}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\psi(x) \\
&= \bar{\psi}(x)\frac{i}{2}[\gamma_\mu, \gamma_\nu]\psi(x) = \bar{\psi}(x)\sigma_{\mu\nu}\psi(x)
\end{aligned} \tag{314}$$

Thus,

$$\bar{\psi}(x)\sigma^{\mu\nu}\psi(x) \xrightarrow{\text{parity}} \bar{\psi}(x)\sigma_{\mu\nu}\psi(x) \tag{315}$$

From Eq.(313), we can say

$$\sigma^{\mu\nu} \xrightarrow{\text{parity}} \sigma_{\mu\nu} \qquad \text{or equivalently} \qquad S^{-1}\sigma^{\mu\nu}S = \sigma_{\mu\nu} \tag{316}$$

1.19 Charge Conjugation

Charge conjugation means we exchange a particle with an antiparticle or vice versa. Let's state relevant statements:

1. Charge conjugation means exchanging a particle with an antiparticle or vice versa. (Definition of Charge Conjugation)

2. An antiparticle means the corresponding particle with an opposite charge traveling forward in time (Eq.(100)).

3. Thus, what the charge conjugation operation does to a physical system is to replace a particle with charge q with an antiparticle with charge $-q$.

Because a particle of charge q and its antiparticle of charge $-q$ behave differently in an electromagnetic field, the first mathematical treatment is to use the canonical momentum.

$$\text{For particle:} \quad p_\mu \to p_\mu - qA_\mu \qquad\qquad p^\mu \to p^\mu - qA^\mu \tag{317}$$

$$\text{For antiparticle:} \quad p_\mu \to p_\mu + qA_\mu \qquad p^\mu \to p^\mu + qA^\mu \tag{318}$$

By using

$$p_\mu = (E, -\vec{p}) \to i\partial_\mu \qquad p^\mu = (E, \vec{p}) \to i\partial^\mu \tag{319}$$

Eq.(317) and Eq.(318) become

$$\text{For particle:} \quad \partial_\mu \to \partial_\mu + iqA_\mu \qquad \partial^\mu \to \partial^\mu + iqA^\mu \tag{320}$$

$$\text{For antiparticle:} \quad \partial_\mu \to \partial_\mu - iqA_\mu \qquad \partial^\mu \to \partial^\mu - iqA^\mu \tag{321}$$

Let us assume that a particle obeys the Klein-Gordon equation, which can be written as:

$$\begin{aligned}\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\psi(t,\vec{x}) = 0 &\to \left(\partial_0\partial^0 + \partial_k\partial^k + m^2\right)\psi(t,\vec{x}) = 0 \\ &\to \left(\partial_\mu\partial^\mu + m^2\right)\psi(t,\vec{x}) = 0\end{aligned}$$

By taking the electromagnetic interaction into account:

$$\begin{aligned}&\left(\partial_\mu\partial^\mu + m^2\right)\psi(t,\vec{x}) = 0 \\ &\xrightarrow{\partial_\mu \to \partial_\mu + iqA_\mu \text{ and } \partial^\mu \to \partial^\mu + iqA^\mu} \left[\left(\partial_\mu + iqA_\mu\right)\left(\partial^\mu + iqA^\mu\right) + m^2\right]\psi(t,\vec{x}) = 0\end{aligned} \tag{322}$$

We know that charge conjugation, by its definition, inverses the sign of q. Then, by taking the Hermitian conjugate of Eq.(322), we can inverse the sign of q.

$$\left[\left(\partial_\mu - iqA_\mu\right)\left(\partial^\mu - iqA^\mu\right) + m^2\right]\psi^\dagger(t,\vec{x}) = 0 \tag{323}$$

Thus, the solution of the charge-conjugated Klein-Gordon equation is $\psi^\dagger(t,\vec{x})$, which corresponds to the antiparticle.

1.19.1 Charge-Conjugated $\psi(x)$

What if a particle obeys the Dirac equation? First of all, the Dirac equation is:

$$(\gamma^\mu i\partial_\mu - m)\psi(x) = 0$$

By replacing ∂_μ with $\partial_\mu + iqA_\mu$ to include the electromagnetic interaction, the Dirac equation becomes:

$$\begin{aligned}&[\gamma^\mu i(\partial_\mu + iqA_\mu) - m]\psi(x) = 0 \\ &\to [\gamma^\mu(i\partial_\mu - qA_\mu) - m]\psi(x) = 0\end{aligned} \tag{324}$$

So, with $q \to -q$, we want to find $\psi^c(x)$ such that

$$[\gamma^\mu(i\partial_\mu + qA_\mu) - m]\psi^c(x) = 0 \tag{325}$$

where $\psi^c(x)$ is the solution of the charge-conjugated Dirac equation. Let's start with taking the Hermitian conjugate on Eq.(324).

$$\psi^\dagger(x)[\gamma^{\mu\dagger}(-i\overleftarrow{\partial_\mu} - qA_\mu) - m] = 0 \tag{326}$$

By multiplying γ^0 from the right and using $\gamma^0\gamma^0 = I$, Eq.(326) becomes:

$$\psi^\dagger(x)\gamma^0\gamma^0[\gamma^{\mu\dagger}(-i\overleftarrow{\partial_\mu} - qA_\mu) - m]\gamma^0 = 0$$

$$\psi^\dagger(x)\gamma^0[\gamma^0\gamma^{\mu\dagger}\gamma^0(-i\overleftarrow{\partial_\mu} - qA_\mu) - \gamma^0 m\gamma^0] = 0 \tag{327}$$

By using Eq.(115(iv)), Eq.(327) is:

$$\bar{\psi}[-\gamma^\mu(i\overleftarrow{\partial_\mu} + qA_\mu) - m] = 0 \tag{328}$$

By taking the transpose of Eq.(328),

$$[-\gamma^{\mu T}(i\partial_\mu + qA_\mu) - m]\bar{\psi}^T = 0 \tag{329}$$

Let us introduce an operator C. By multiplying C from the left on Eq.(329) and using $C^{-1}C = I$, Eq.(329) becomes:

$$C[-\gamma^{\mu T}(i\partial_\mu + qA_\mu) - m]C^{-1}C\bar{\psi}^T = 0$$

$$[-C\gamma^{\mu T}C^{-1}(i\partial_\mu + qA_\mu) - m]C\bar{\psi}^T = 0 \tag{330}$$

If we can find C such that $C\gamma^{\mu T}C^{-1} = -\gamma^\mu$, we can claim that $C\bar{\psi}^T$ is a solution of the charge-conjugated Dirac equation, by comparing with Eq.(325).

1.19.2 C exists!

Assuming such C exists (and it does),

$$\psi^c(x) = C\bar{\psi}^T \tag{331}$$

So what we have to find is C such that

$$C\gamma^{\mu T}C^{-1} = -\gamma^\mu \tag{332}$$

However, before doing so, let us calculate the transpose of each Pauli matrix.

$$\sigma_1^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1 \tag{333}$$

$$\sigma_2^T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = -\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -\sigma_2 \tag{334}$$

$$\sigma_3^T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3 \tag{335}$$

Let us calculate $\gamma^{\mu T}$ for each μ.

1. $\mu = 0$

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \to (\gamma^0)^T = \gamma^0 \tag{336}$$

2. $\mu = 1$

$$\begin{aligned} \gamma^1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} \to (\gamma^1)^T &= \begin{pmatrix} 0 & -\sigma_1^T \\ \sigma_1^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\sigma_1 \\ \sigma_1 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} = -\gamma^1 \\ \to (\gamma^1)^T &= -\gamma^1 \end{aligned} \tag{337}$$

3. $\mu = 2$

$$\begin{aligned} \gamma^2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \to (\gamma^2)^T = \begin{pmatrix} 0 & -\sigma_2^T \\ \sigma_2^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = \gamma^2 \\ \to (\gamma^2)^T = \gamma^2 \end{aligned} \tag{338}$$

4. $\mu = 3$

$$\begin{aligned} \gamma^3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} \to (\gamma^3)^T &= \begin{pmatrix} 0 & -\sigma_3^T \\ \sigma_3^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\sigma_3 \\ \sigma_3 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} = -\gamma^3 \\ \to (\gamma^3)^T &= -\gamma^3 \end{aligned} \tag{339}$$

where we used Eq.(333), (334), and (335).

We can summarize Eq.(336), (337), (338), and (339) as follows:

$$(\gamma^\mu)^T = \gamma^\mu \qquad \mu = 0, 2 \qquad\qquad (\gamma^\mu)^T = -\gamma^\mu \qquad \mu = 1, 3 \tag{340}$$

Let us go back to Eq.(332).

$$C\gamma^{\mu T} C^{-1} = -\gamma^\mu \to C\gamma^{\mu T} = -\gamma^\mu C \tag{341}$$

With the help of Eq.(340), Eq.(341) can be written as:

$$C\gamma^\mu = -\gamma^\mu C \qquad \mu = 0, 2 \qquad\qquad C\gamma^\mu = \gamma^\mu C \qquad \mu = 1, 3 \tag{342}$$

We also know the anti-commutation relation between γ^i and γ^j (Eq.(108.iii)):

$$\{\gamma^i, \gamma^j\} = 0 \to \gamma^i\gamma^j = -\gamma^j\gamma^i$$

which means that switching γ^i and γ^j produces a minus sign.

So, C can be defined as the following:

$$C = i\gamma^2\gamma^0 = i\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} \tag{343}$$

so that, when switched with γ^0 or γ^2, we use the anti-commutation relation of the gamma matrices *one* time, and thus pick up *one* minus sign – on the other hand, when switched with γ^1 or γ^3, we use the anti-commutation relation of the gamma matrices *two* times, and thus pick up *two* minus signs, which is a positive sign. C can be explicitly written as:

$$C = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{344}$$

$$C^2 = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix}\begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} = \begin{pmatrix} -\sigma_2^2 & 0 \\ 0 & -\sigma_2^2 \end{pmatrix} = -I \tag{345}$$

where we used $\sigma_i^2 = I$.

1.19.3 Specific Form of $\psi^c(x)$

From Eq.(344) and (345), we see that

$$C = C^* = -C^\dagger = -C^T = -C^{-1} \tag{346}$$

Thus the charge-conjugated solution to the Dirac equation is:

$$\begin{aligned} \psi^c(x) = C\bar{\psi}^T(x) &= (i\gamma^2\gamma^0)(\psi^\dagger(x)\gamma^0)^T \\ &= i\gamma^2\gamma^0(\gamma^0)^T\psi^*(x) = i\gamma^2\gamma^0\gamma^0\psi^*(x) = i\gamma^2\psi^*(x) \\ &\to \psi^c(x) = i\gamma^2\psi^*(x) \end{aligned} \tag{347}$$

$$\begin{aligned} \bar{\psi^c}(x) = [\psi^c(x)]^\dagger\gamma^0 &= [i\gamma^2\psi^*(x)]^\dagger\gamma^0 = [\psi^*(x)]^\dagger(-i\gamma^{2\dagger}\gamma^0) \\ &= [\psi(x)]^T(i\gamma^2\gamma^0) = \psi^T(x)C = -\psi^T(x)C^{-1} \\ &\to \bar{\psi^c}(x) = -\psi^T(x)C^{-1} \end{aligned} \tag{348}$$

where we used Eq.(112(iii)).

1.19.4 Relation with $v_r(p)$

Let's check if this discussion is actually valid. We have seen that Eq.(219) and Eq.(227) describe a particle and its antiparticle, respectively. We want to see if the charge-conjugated Eq.(219) produces Eq.(227). That is,

$$\begin{aligned}(u_r)^c &= C\bar{u}_r^T = i\gamma^2\gamma^0(u_r^\dagger\gamma^0)^T = i\gamma^2\gamma^0(\gamma^0)^T(u_r^\dagger)^T = i\gamma^2 u_r^* \\ &= \begin{pmatrix} 0 & i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} N\sqrt{m} \begin{pmatrix} [e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \\ [e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \end{pmatrix} = N\sqrt{m} \begin{pmatrix} i\sigma_2[e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \\ -i\sigma_2[e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \end{pmatrix}\end{aligned} \tag{349}$$

We have to calculate $\sigma_2[e^{\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*$. Here are the Pauli matrices (I keep forgetting them, so I again write them here):

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Notice that only σ_2 changes its sign and σ_1 and σ_3 remain intact under taking the complex conjugate. Also, remember

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}$$

$$\sigma_2[e^{\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^* = \sigma_2\left[1 + \frac{1}{1!}\left(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^1 + \frac{1}{2!}\left(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^2 + \cdots\right]^* \tag{350}$$

Let us calculate the second term:

$$\begin{aligned}\sigma_2\frac{1}{1!}\left(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^* &= \frac{1}{1!}\left(\pm\frac{1}{2}\right)\sigma_2(\vec{\sigma}\cdot\vec{\eta})^* = \frac{1}{1!}\left(\pm\frac{1}{2}\right)\sigma_2(\sigma_1\eta_1 + \sigma_2\eta_2 + \sigma_3\eta_3)^* \\ &= \frac{1}{1!}\left(\pm\frac{1}{2}\right)\sigma_2(\sigma_1\eta_1 - \sigma_2\eta_2 + \sigma_3\eta_3) \\ &= \frac{1}{1!}\left(\pm\frac{1}{2}\right)(\sigma_2\sigma_1\eta_1 - \sigma_2\sigma_2\eta_2 + \sigma_2\sigma_3\eta_3) \\ &= \frac{1}{1!}\left(\pm\frac{1}{2}\right)(-\sigma_1\sigma_2\eta_1 - \sigma_2\sigma_2\eta_2 - \sigma_3\sigma_2\eta_3) \\ &= \frac{1}{1!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)\sigma_2\end{aligned} \tag{351}$$

Therefore,

$$\sigma_2\frac{1}{1!}\left(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^* = \frac{1}{1!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)\sigma_2 \tag{352}$$

This means that if we switch σ_2 and $(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta})^*$, $(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta})^*$ flips its sign inside the parenthesis and "uncomplex-conjugated." Then we can deduce the following:

$$\sigma_2\frac{1}{n!}\left[\left(\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^n\right]^* = \frac{1}{1!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^n\sigma_2 \tag{353}$$

Hence, Eq.(350) can be written as:

$$\begin{aligned}
\sigma_2[e^{\pm\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^* &= \sigma_2\left[1+\frac{1}{1!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^1+\frac{1}{2!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^2+\cdots\right]^* \\
&= \left[1+\frac{1}{1!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^1+\frac{1}{2!}\left(\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}\right)^2+\cdots\right]\sigma_2 \\
&= e^{\mp\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\sigma_2
\end{aligned} \tag{354}$$

Thus Eq.(349) can be written as:

$$\begin{aligned}
(u_r)^c &= N\sqrt{m}\begin{pmatrix} i\sigma_2[e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \\ -i\sigma_2[e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\xi_r^* \end{pmatrix} = N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}i\sigma_2\xi^* \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}i\sigma_2\xi^* \end{pmatrix} \\
&= N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \\ -e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\eta_r \end{pmatrix} = v_r
\end{aligned} \tag{355}$$

which is exactly Eq.(227). Now you see why we suddenly introduced Eq.(229) before.

We can also show $(v_r)^c = u_r$.

$$\begin{aligned}
(v_r)^c &= C\bar{v}_r^T = i\gamma^2\gamma^0(v_r^\dagger\gamma^0)^T = i\gamma^2\gamma^0(\gamma^0)^T(v_r^\dagger)^T = i\gamma^2 v_r^* \\
&= \begin{pmatrix} 0 & i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} N\sqrt{m}\begin{pmatrix} [e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\eta_r^* \\ -[e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\eta_r^* \end{pmatrix} = N\sqrt{m}\begin{pmatrix} -i\sigma_2[e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\eta_r^* \\ -i\sigma_2[e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]^*\eta_r^* \end{pmatrix} \\
&= N\sqrt{m}\begin{pmatrix} -i[e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]\sigma_2\eta_r^* \\ -i[e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}]\sigma_2\eta_r^* \end{pmatrix} = N\sqrt{m}\begin{pmatrix} e^{-\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \\ e^{\frac{1}{2}\vec{\sigma}\cdot\vec{\eta}}\xi_r \end{pmatrix} = u_r
\end{aligned} \tag{356}$$

where we used $\xi_r = -i\sigma_2\eta_r^*$.

2 Lagrangian in QFT and Field Quantization

Lagrangian is a packet which contains all the information about the particle under consideration. It is often helpful to know the equation of motion first and come up with the Lagrangian. For example, people knew Newton's second law and then defined the Lagrangian to be $L = T - U$, such that the Euler-Lagrangian equation gives Newton's second law.

Chapter 1 introduced equations of motion (Klein-Gordon and Dirac equation) and investigated their solutions. So we can ask how the Lagrangians have to be defined such that the Euler-Lagrange equation produces the Klein-Gordon or Dirac equation. Before doing so, let us remind ourselves of Lagrangian formalism in classical mechanics.

2.1 Lagrangian in Classical Mechanics

Action principle states that the equation of motion of a particle is given such that the following integral is minimized:

$$S[q] = \int_{q_i}^{q_f} L(q,\dot{q})dt \qquad\qquad L = T - U \tag{357}$$

where $\dot{q} = \frac{dq}{dt}$.

Eq.(357) is a function which takes a function $q = q(t)$ as an input and produces a number as an output. What the action principle tells us is that the actual path that a particle takes, out of all possible paths, is given by the path which minimizes the integral.

Let us see *how* this principle helps us to find an equation of motion. The idea is as follows: for the path the integral is already minimized, if we deviate from the path by a small amount δq *with the endpoints fixed*, the small change of the integral, δS, is zero up to the first order. It is like the first derivative is zero at extrema. So let us calculate δS.

$$\begin{aligned}
\delta S &= \delta \int_{t_i}^{t_f} L(q,\dot{q})dt = \int_{t_i}^{t_f} \delta L(q,\dot{q})dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} \right) dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q}\delta q \right) dt + \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial \dot{q}} \frac{d(\delta q)}{dt} \right) dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q}\delta q \right) dt + \left[\frac{\partial L}{\partial \dot{q}}\delta q \right]_{t_i}^{t_f} - \int_{t_i}^{t_f} \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\delta q \right) dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) \delta q dt = 0
\end{aligned} \tag{358}$$

where we used $\delta q(t_i) = \delta q(t_f) = 0$, which is equivalent of saying the endpoints are fixed.

Thus $\delta S = 0$ for *any* δq. This implies that the integrand must be zero. Hence we have:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \tag{359}$$

For example, if $L = \frac{1}{2}m\dot{q}^2 - U(q)$, Eq.(359) produces:

$$-\frac{\partial U}{\partial q} = \frac{d}{dt}(m\dot{q}) = m\ddot{q} \tag{360}$$

which is Newton's second law.

2.2 Euler-Lagrange Equation in QFT

We want to follow an analogous version of $L = T - V$. That is, we want to find the Lagrangians whose Euler-Lagrange equations produce the Klein-Gordon, Dirac, or other equations of motion. To do that, we expand our concept about $L(q, \dot{q})$. Let us just repeat the Klein-Gordon and Dirac equation:

$$\begin{aligned}&\text{Klein-Gordon Equation: } \left(\partial_\mu \partial^\mu + m^2\right) \psi(t, \vec{x}) = 0 \\ &\text{Dirac Equation: } (\gamma^\mu i \partial_\mu - m)\psi(x) = 0\end{aligned} \tag{361}$$

Then we compare Eq.(360) and Eq.(361). First of all, we have $\psi(x)$ instead of q. Also, we have spatial derivatives of $\psi(x)$, which are not present in Eq.(359), as they should be because Eq.(360) cannot have spatial derivatives of q, which is a spatial coordinate itself. To sum up,

$$\psi(x) \to q \qquad\qquad \dot{q} \to \partial_\mu \psi(x) \tag{362}$$

Thus, it seems that $\partial_t \to \partial_\mu$. Then we can say:

$$\int dt \to \int d^4x = \int dt \int d^3x \tag{363}$$

If we act the integrals on a Lagrangian,

$$S = \int L dt \to S = \int \mathscr{L} d^4x = \int dt \int \mathscr{L} d^3x = \int L dt \tag{364}$$

Because spatial integration is involved, it is natural to introduce Lagrangian density, $\mathscr{L}$, whose spatial integration produces Lagrangian. Once we integrate the Lagrangian over time, we get the action integral.

$$\int \mathscr{L} d^3x = L \tag{365}$$

Thus, with Eq.(362), a Lagrangian density can be expressed as $\mathscr{L}(\psi(x), \partial_\mu \psi(x))$.

We can derive the Euler-Lagrange equation for $\mathscr{L}(\psi(x), \partial_\mu \psi(x))$ by following the similar procedure as Eq.(358):

$$\begin{aligned}\delta S &= \delta \int \mathscr{L}(\psi(x), \partial_\mu \psi(x)) d^4x = \int \delta \mathscr{L}(\psi(x), \partial_\mu \psi(x)) d^4x \\ &= \int \left(\frac{\partial \mathscr{L}}{\partial \psi} \delta\psi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \delta(\partial_\mu \psi) \right) d^4x \\ &= \int \left(\frac{\partial \mathscr{L}}{\partial \psi} \delta\psi \right) d^4x + \int \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \partial_\mu (\delta\psi) \right) d^4x \\ &= \int \left(\frac{\partial \mathscr{L}}{\partial \psi} \delta\psi \right) d^4x + \left[\frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \delta\psi \right]_{\text{boundary}} - \int \left(\partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \delta\psi \right) d^4x \\ &= \int \left(\frac{\partial \mathscr{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} \right) \delta\psi d^4x = 0\end{aligned} \tag{366}$$

Therefore, an equation of motion is given by

$$\frac{\partial \mathscr{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)} = 0 \tag{367}$$

2.3 Lagrangians in QFT

Let me just throw the known Lagrangians to you.

$$\mathscr{L}_{KG,r} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2\phi^2) \qquad \text{Real field} \tag{368}$$

$$\mathscr{L}_{KG,c} = \partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi \qquad \text{Complex field} \tag{369}$$

$$\mathscr{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \qquad \text{Dirac field} \tag{370}$$

$$\mathscr{L}_{EM} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - A_\mu J^\mu \qquad \text{Electromagnetic field} \tag{371}$$

where

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \qquad\qquad F^{\nu\mu} = -F^{\mu\nu} \tag{372}$$

1. Real Field

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} &= -m^2\phi - \frac{1}{2}\partial_\mu \frac{\partial(\partial_\alpha \phi \partial^\alpha \phi)}{\partial(\partial_\mu \phi)} \\
&= -m^2\phi - \frac{1}{2}\partial_\mu \left(\frac{\partial(\partial_a \phi)}{\partial(\partial_\mu \phi)} \partial^\alpha \phi + \partial_\alpha \phi \frac{\partial(\partial^\alpha \phi)}{\partial(\partial_\mu \phi)} \right) \\
&= -m^2\phi - \frac{1}{2}\partial_\mu \left(\delta^\mu_\alpha \partial^\alpha \phi + \partial_\alpha \phi \frac{\partial(\partial_\beta \phi)}{\partial(\partial_\mu \phi)} g^{\alpha\beta} \right) \\
&= -m^2\phi - \frac{1}{2}\partial_\mu \left(\partial^\mu \phi + \partial_\alpha \phi \delta^\mu_\beta g^{\alpha\beta} \right) \\
&= -m^2\phi - \frac{1}{2}\partial_\mu \left(\partial^\mu \phi + \partial^\mu \phi \right) \\
&= -m^2\phi - \partial_\mu \partial^\mu \phi = 0
\end{aligned} \tag{373}$$

which is the Klein-Gordon equation.

2. Complex Field

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \varphi^\dagger} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \varphi^\dagger)} &= -m^2\varphi - \partial_\mu \frac{\partial(\partial_\alpha \varphi^\dagger \partial^\alpha \varphi)}{\partial(\partial_\mu \varphi^\dagger)} \\
&= -m^2\varphi - \partial_\mu \left(\frac{\partial(\partial_a \varphi^\dagger)}{\partial(\partial_\mu \varphi^\dagger)} \partial^\alpha \varphi + \partial_\alpha \varphi^\dagger \frac{\partial(\partial^\alpha \varphi)}{\partial(\partial_\mu \varphi^\dagger)} \right) \\
&= -m^2\varphi - \partial_\mu \left(\delta^\mu_\alpha \partial^\alpha \varphi + 0 \right) \\
&= -m^2\varphi - \partial_\mu \left(\partial^\mu \varphi \right) \\
&= -m^2\varphi - \partial_\mu \left(\partial^\mu \varphi \right) \\
&= -m^2\varphi - \partial_\mu \partial^\mu \varphi = 0
\end{aligned} \tag{374}$$

which is the Klein-Gordon equation.

3. Dirac Field

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial \bar{\psi}} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \bar{\psi})} &= (i\gamma^\mu \partial_\mu - m)\psi - 0 \\ &= (i\gamma^\mu \partial_\mu - m)\psi = 0\end{aligned} \tag{375}$$

which is the Dirac equation.

4. Electromagnetic Field
Before calculating the Euler-Lagrange equation, we expand the following:

$$\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)} = \frac{\partial(\partial_\alpha A_\beta - \partial_\beta A_\alpha)}{\partial(\partial_\mu A_\nu)} = \delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha \tag{376}$$

$$\frac{\partial F^{\alpha\beta}}{\partial(\partial_\mu A_\nu)} = \frac{\partial F_{\sigma\rho}}{\partial(\partial_\mu A_\nu)} g^{\alpha\sigma} g^{\beta\rho} = \left(\delta^\mu_\sigma \delta^\nu_\rho - \delta^\mu_\rho \delta^\nu_\sigma\right) g^{\alpha\sigma} g^{\beta\rho} = g^{\alpha\mu} g^{\beta\nu} - g^{\alpha\nu} g^{\beta\mu} \tag{377}$$

Let us proceed to compute the Euler-Lagrange equation.

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial A_\nu} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{\partial A_\mu}{\partial A_\nu} J^\mu + \frac{1}{4}\partial_\mu \frac{\partial(F_{\alpha\beta}F^{\alpha\beta})}{\partial(\partial_\mu A_\nu)} \\ &= -\delta^\nu_\mu J^\mu + \frac{1}{4}\partial_\mu \left(\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)} F^{\alpha\beta} + F_{\alpha\beta}\frac{\partial F^{\alpha\beta}}{\partial(\partial_\mu A_\nu)}\right) \\ &= -J^\nu + \frac{1}{4}\partial_\mu \left[\left(\delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha\right) F^{\alpha\beta} + F_{\alpha\beta}\left(g^{\alpha\mu} g^{\beta\nu} - g^{\alpha\nu} g^{\beta\mu}\right)\right] \\ &= -J^\nu + \frac{1}{4}\partial_\mu \left(F^{\mu\nu} - F^{\nu\mu} + F^{\mu\nu} - F^{\nu\mu}\right) \\ &= -J^\nu + \partial_\mu F^{\mu\nu} = 0\end{aligned} \tag{378}$$

where we used $F^{\nu\mu} = -F^{\mu\nu}$ and the last line is Maxwell's equations.

2.4 Hamiltonian in Classical Mechanics

While the Euler-Lagrange equation of a Lagrangian gives you the equation of motion, the Hamiltonian equation derived from the Hamiltonian also gives you the equation of motion. Lagrangian is expressed by $L(q, \dot{q})$, Hamiltonian is given by $H(p, q)$ where p is momentum. Given that a Lagrangian is given by:

$$L = \frac{1}{2}m\dot{q}^2 - V(q)$$

the momentum p can be defined as:

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \tag{379}$$

We can use the chain rule on $L(p, \dot{p})$:

$$\delta L = \frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} = \frac{\partial L}{\partial q}\delta q + p\delta \dot{q} \tag{380}$$

By using

$$\delta(p\dot{q}) = (\delta p)\dot{q} + p(\delta \dot{q}) \to p(\delta \dot{q}) = \delta(p\dot{q}) - (\delta p)\dot{q}$$

we can proceed expanding Eq.(380),

$$\delta L = \frac{\partial L}{\partial q}\delta q + p\delta \dot{q} = \frac{\partial L}{\partial q}\delta q + \delta(p\dot{q}) - (\delta p)\dot{q}$$

By moving around the terms, we have:

$$\delta(p\dot{q} - L) = -\frac{\partial L}{\partial q}\delta q + (\delta p)\dot{q} = -\dot{p}\delta q + \dot{q}\delta p \tag{381}$$

where we used the Euler-Lagrange equation on the first term. In fact, Hamiltonian corresponds to the total energy, which is

$$H(p, q) = \frac{1}{2m}p^2 + V(q) = p\dot{q} - L \tag{382}$$

Then we can use Eq.(382) in Eq.(381):

$$\delta H(p, q) = -\dot{p}\delta q + \dot{q}\delta p = \frac{\partial H}{\partial q}\delta q + \frac{\partial H}{\partial p}\delta p \tag{383}$$

where we used the chain rule on the last equality. Then from Eq.(383), we have:

$$\frac{\partial H}{\partial q} = -\dot{p} \qquad\qquad \frac{\partial H}{\partial p} = \dot{q} \tag{384}$$

With $H(p, q) = \frac{1}{2m}p^2 + V(q)$, Eq.(384) becomes:

$$\frac{\partial V}{\partial q} = -\dot{p} \qquad\qquad \frac{p}{m} = \dot{q} \tag{385}$$

which is Newton's second law.

2.5 Hamiltonian Formalism in QFT

Just as momentum was defined as the partial derivative of L with respect to $\dot{q}$ (Eq.(379)), the momentum field $\pi(x)$ is given by:

$$p = \frac{\partial L}{\partial \dot{q}} \to \pi(x) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}} \tag{386}$$

where we used $q \to \phi(x)$.

Then, Eq.(382) hints how the Hamiltonian density is given:

$$H(p, q) = p\dot{q} - L \to \mathscr{H}(x) = \pi(x)\dot{\phi}(x) - \mathscr{L}(\phi, \partial_\mu \phi) \tag{387}$$

With this definition of Hamiltonian density, we can derive the following theorem.

The Hamiltonian is constant if the Lagrangian density is constant in time.

$$\frac{dH}{dt} = \frac{d}{dt}\int \mathscr{H} d^3x = 0 \quad \text{if } \frac{\partial \mathscr{L}}{\partial t} = 0 \tag{388}$$

proof)

$$\begin{aligned}
\frac{dH}{dt} &= \frac{d}{dt}\int \mathscr{H} d^3x = \int \dot{\mathscr{H}} d^3x = \int \frac{d}{dt}\Big(\pi(x)\dot{\phi}(x) - \mathscr{L}(\phi, \partial_\mu\phi)\Big) d^3x \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) + \pi(x)\ddot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \frac{\partial \mathscr{L}}{\partial \phi}\frac{\partial \phi}{\partial t} + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\frac{\partial(\partial_\mu\phi)}{\partial t}\right)\right] d^3x \quad \text{Chain rule on } \frac{d\mathscr{L}}{dt} \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) + \pi(x)\ddot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \frac{\partial \mathscr{L}}{\partial \phi}\frac{\partial \phi}{\partial t} + \frac{\partial \mathscr{L}}{\partial(\partial_t \phi)}\frac{\partial(\partial_t\phi)}{\partial t} + \frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\frac{\partial(\partial_k\phi)}{\partial t}\right)\right] d^3x \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) + \pi(x)\ddot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \frac{\partial \mathscr{L}}{\partial \phi}\dot{\phi} + \pi(x)\ddot{\phi} + \frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\partial_k\dot{\phi}\right)\right] d^3x \quad \text{Used } \pi = \frac{\partial \mathscr{L}}{\partial(\partial_t\phi)} \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \frac{\partial \mathscr{L}}{\partial \phi}\dot{\phi} + \frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\partial_k\dot{\phi}\right)\right] d^3x \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \partial_\mu\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\dot{\phi} - \partial_k\frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\dot{\phi}\right)\right] d^3x \\
&\quad \text{Used } \frac{\partial \mathscr{L}}{\partial \phi} = \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)} \text{ and } \int \frac{\partial \mathscr{L}}{\partial(\partial_k\phi)}\partial_k\dot{\phi} d^3x = \left[\frac{\partial \mathscr{L}}{\partial(\partial_k\phi)}\dot{\phi}\right]_{\text{Boundary}} - \int \partial_k \frac{\partial \mathscr{L}}{\partial(\partial_k\phi)}\dot{\phi} d^3x \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \partial_0\frac{\partial \mathscr{L}}{\partial(\partial_0 \phi)}\dot{\phi}\right)\right] d^3x \\
&= \int \left[\dot{\pi}(x)\dot{\phi}(x) - \left(\frac{\partial \mathscr{L}}{\partial t} + \dot{\pi}(x)\dot{\phi}\right)\right] d^3x \\
&= -\int \frac{\partial \mathscr{L}}{\partial t} d^3x = 0 \quad \text{if } \frac{\partial \mathscr{L}}{\partial t} = 0
\end{aligned} \tag{389}$$

2.6 Noether's Theorem

Eq.(389) implies that the energy of a field is conserved if the Lagrangian density does not explicitly depend on time. Then we ask: Are there more conserved quantities other than energy? One very interesting thing about physics is that if a system is invariant under a certain transformation, there is a conserved quantity. Consider an infinitesimal change of a field ϕ. Then the Lagrangian $\mathscr{L}$

also changes by an infinitesimal amount.

$$\begin{aligned}
\delta\mathscr{L} &= \frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\delta(\partial_\mu\phi_r) = \frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\mu(\delta\phi_r) \\
&= \frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \partial_\mu\left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}(\delta\phi_r)\right] - \partial_\mu\left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\right](\delta\phi_r) \qquad \text{Product Rule} \\
&= \left(\frac{\partial\mathscr{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\right)\delta\phi_r + \partial_\mu\left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}(\delta\phi_r)\right] \\
&= \partial_\mu\left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}(\delta\phi_r)\right] \qquad \text{Used the Euler-Lagrange Equation}
\end{aligned} \tag{390}$$

Thus, if the Lagrangian is invariant under some continuous transformation of ϕ, the following must be true:

$$\delta\mathscr{L} = \partial_\mu\left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}(\delta\phi_r)\right] = 0 \tag{391}$$

So we define the following:

$$J^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}(\delta\phi_r) \tag{392}$$

Then, by Eq.(391), $\partial_\mu J^\mu = 0$. So we call Eq.(392) a conserved current. Then we have the following property.

$$\frac{dQ}{dt} = 0, \qquad Q = \int J^0 d^3x \tag{393}$$

proof

$$\begin{aligned}
\frac{dQ}{dt} &= \frac{d}{dt}\int J^0 d^3x = \int \partial_0 J^0 d^3x \\
&= -\int \nabla\cdot\vec{J}d^3x = -\int_{\text{boundary}} \vec{J}\cdot\hat{n}da = 0
\end{aligned} \tag{394}$$

where we used $\partial_0 J^0 = -\partial_k J^k$ and the divergence theorem. $\hat{n}$ is a vector normal to the surface at infinitely far away.

2.7 Conserved Current Example 1

Let us work on an example to calculate J^μ. Consider Eq.(369):

$$\mathscr{L}_{KG,c} = \partial_\mu\varphi^\dagger\partial^\mu\varphi - m^2\varphi^\dagger\varphi$$

We see that the Lagrangian is invariant under the following substitution:

$$\varphi \to e^{-iq\alpha}\varphi \qquad\qquad \varphi^\dagger \to \varphi^\dagger e^{iq\alpha} \tag{395}$$

where α is a parameter which does not depend on space-time coordinates x^μ. By Eq.(395),

$$\delta\varphi = -iq\delta\alpha\varphi \qquad\qquad \delta\varphi^\dagger = iq\delta\alpha\varphi^\dagger \tag{396}$$

Now we can calculate J^μ

$$\begin{aligned} J^\mu &= \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi_r)}(\delta\varphi_r) = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi)}\delta\varphi + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi^\dagger)}\delta\varphi^\dagger \\ &= \partial^\mu\varphi^\dagger(-iq\delta\alpha\varphi) + \partial^\mu\varphi(iq\delta\alpha\varphi^\dagger) \\ &= iq(\varphi^\dagger\partial^\mu\varphi - (\partial^\mu\varphi^\dagger)\varphi)\delta\alpha \end{aligned} \tag{397}$$

2.8 Conserved Current Example 2

Consider Eq.(370) to calculate J^μ:

$$\mathscr{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi$$

We see that the Lagrangian is invariant under the following substitution:

$$\psi \to e^{-i\alpha}\psi \qquad\qquad \bar{\psi} \to \bar{\psi}e^{i\alpha} \tag{398}$$

where α is a parameter which does not depend on space-time coordinates x^μ. By Eq.(398),

$$\delta\psi = -i\delta\alpha\psi \qquad\qquad \delta\bar{\psi} = i\delta\alpha\bar{\psi} \tag{399}$$

Now we can calculate J^μ

$$\begin{aligned} J^\mu &= \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi_r)}(\delta\varphi_r) = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\psi)}\delta\psi + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\bar{\psi})}\delta\bar{\psi} \\ &= \bar{\psi}i\gamma^\mu(-i\delta\alpha) \\ &= \bar{\psi}\gamma^\mu\psi\delta\alpha \end{aligned} \tag{400}$$

Eq.(400) looks familiar! Eq.(121) has it.

2.9 Energy-Momentum Tensor

The Noether's theorem demonstrates that we can extract the conserved current when the Lagrangian is invariant under the phase transformation ($\phi \to e^{-i(\text{phase})}\phi$). However, we can consider a case where a Lagrangian is invariant under an infinitesimal translation:

$$x'^\mu = x^\mu + \delta x^\mu \tag{401}$$

Then the Lagrangian changes

$$\mathscr{L}(\phi(x), \partial_\mu\phi(x)) \to \mathscr{L}(\phi(x+\delta x), \partial_\mu\phi(x+\delta x)) = \mathscr{L}(\phi+\delta\phi, \partial_\mu\phi+\delta(\partial_\mu\phi)) \tag{402}$$

Let us calculate the infinitesimal change of ϕ so that we can calculate Eq.(402).

$$\delta\phi_r = \phi_r(x^\mu + \delta x^\mu) - \phi_r(x^\mu) = \partial_\mu\phi_r\delta x^\mu \tag{403}$$

$$\delta(\partial_\mu\phi_r) = \partial_\mu\phi_r(x^\nu + \delta x^\nu) - \partial_\mu\phi_r(x^\nu) = \partial_\nu\partial_\mu\phi_r\delta x^\nu \tag{404}$$

Then the change of the Lagrangian can be written in two ways:

$$\delta\mathscr{L} = \partial_\mu\mathscr{L}\delta x^\mu = \frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\delta(\partial_\mu\phi_r) \tag{405}$$

Using the fact that $\delta S = 0$,

$$0 = \delta S = \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\delta(\partial_\mu\phi_r) - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x \tag{406}$$

Inserting Eq.(403) and Eq.(404) into Eq.(406),

$$\begin{aligned}
0 = \delta S &= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\delta(\partial_\mu\phi_r) - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x \\
&= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\partial_\mu\phi_r\delta x^\mu + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\partial_\mu\phi_r\delta x^\nu - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x
\end{aligned} \tag{407}$$

We can use the product rule on the second term.

$$\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\partial_\mu\phi_r\delta x^\nu = \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r\delta x^\nu\right) - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r\delta x^\nu \tag{408}$$

We can use Eq.(408) into Eq.(407)

$$\begin{aligned}
0 = \delta S &= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\delta(\partial_\mu\phi_r) - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x \\
&= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\partial_\mu\phi_r\delta x^\mu + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\partial_\mu\phi_r\delta x^\nu - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x \\
&= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r}\partial_\mu\phi_r\delta x^\mu - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r\delta x^\nu + \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r\delta x^\nu\right) - \partial_\mu\mathscr{L}\delta x^\mu\right)d^4x \\
&= \int\left(\frac{\partial\mathscr{L}}{\partial\phi_r} - \partial_\nu\frac{\partial\mathscr{L}}{\partial(\partial_\nu\phi_r)}\right)\partial_\mu\phi_r\delta x^\mu d^4x + \int\partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r - \delta^\mu_\nu\mathscr{L}\right)\delta x^\nu d^4x
\end{aligned} \tag{409}$$

The first integral is zero because the integrand is the Euler-Lagrange equation. So we are left with only the second integral.

$$0 = \delta S = \int\partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r - \delta^\mu_\nu\mathscr{L}\right)\delta x^\nu d^4x \tag{410}$$

Since the integral is zero for an arbitrary δx^ν, the integrand must be zero:

$$\partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial_\nu\phi_r - \delta^\mu_\nu\mathscr{L}\right) = 0 \tag{411}$$

Thus, we define a tensor T^μ_σ called the energy-momentum tensor:

$$\partial_\mu T^\mu_\sigma = 0 \qquad\qquad T^\mu_\sigma = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_r)}\partial_\sigma \phi_r - \delta^\mu_\sigma \mathscr{L} \tag{412}$$

By multiplying $g^{\sigma\nu}$, Eq.(412) can be written as:

$$\partial_\mu T^{\mu\nu} = 0 \qquad\qquad T^{\mu\nu} = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_r)}\partial^\nu \phi_r - g^{\mu\nu} \mathscr{L} \tag{413}$$

Let us take $\mu = \nu = 0$ of $T^{\mu\nu}$:

$$T^{00} = \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_r)}\partial^0 \phi_r - g^{00}\mathscr{L} = \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_r)}\partial^0 \phi_r - \mathscr{L} \tag{414}$$

Eq.(414) agrees with Eq.(387), which defines the Hamiltonian density. Thus, its spatial integral is constant over time (See Eq.(388)).

2.9.1 Example of Energy-Momentum Tensor – Complex Field

Let us work on an example of a complex Klein-Gordon Lagrangian.

$$\mathscr{L}_{KG,c} = \partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi$$

Then, its energy-momentum tensor is given by:

$$\begin{aligned}
T^{\mu\nu} &= \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_r)}\partial^\nu \phi_r - g^{\mu\nu}\mathscr{L} \\
&= \frac{\partial \mathscr{L}}{\partial(\partial_\mu \varphi)}\partial^\nu \varphi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \varphi^\dagger)}\partial^\nu \varphi^\dagger - g^{\mu\nu}\mathscr{L} \\
&= \partial^\mu \varphi^\dagger \partial^\nu \varphi + \partial^\mu \varphi \partial^\nu \varphi^\dagger - g^{\mu\nu}\mathscr{L}
\end{aligned} \tag{415}$$

The Hamiltonian density is given by:

$$\begin{aligned}
\mathscr{H}(x) = T^{00} &= \partial^0 \varphi^\dagger \partial^0 \varphi + \partial^0 \varphi \partial^0 \varphi^\dagger - g^{00}\mathscr{L} \\
&= \partial^0 \varphi^\dagger \partial^0 \varphi + \partial^0 \varphi \partial^0 \varphi^\dagger - (\partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi) \\
&= \partial^0 \varphi \partial^0 \varphi^\dagger - \partial_k \varphi^\dagger \partial^k \varphi + m^2 \varphi^\dagger \varphi \\
&= \partial^0 \varphi^\dagger \partial^0 \varphi + \sum_{k=1}^{3} \partial_k \varphi^\dagger \partial_k \varphi + m^2 \varphi^\dagger \varphi \\
&= \partial^0 \varphi^\dagger \partial^0 \varphi + \nabla \varphi^\dagger \cdot \nabla \varphi + m^2 \varphi^\dagger \varphi
\end{aligned} \tag{416}$$

where we used $\partial^0 \varphi \partial^0 \varphi^\dagger = \partial^0 \varphi^\dagger \partial^0 \varphi$ because φ is a scalar field ($\varphi^\dagger = \varphi^*$). Momentum density is given by:

$$\begin{aligned}
\mathscr{P}_i(x) = \sum_r T^{0i} &= \sum_r \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_r)}\partial^i \phi_r \\
&= \frac{\partial \mathscr{L}}{\partial(\partial_0 \varphi)}\partial^i \varphi + \frac{\partial \mathscr{L}}{\partial(\partial_0 \varphi)}\partial^i \varphi \\
&= \partial^0 \varphi^\dagger \partial^i \varphi + \partial^0 \varphi \partial^i \varphi^\dagger \\
&= -(\partial^0 \varphi^\dagger \partial_i \varphi + \partial_i \varphi^\dagger \partial^0 \varphi)
\end{aligned} \tag{417}$$

We can also calculate the conserved charge density from Eq.(397):

$$\mathscr{Q}(x) = J^0 = iq(\varphi^\dagger \partial^0 \varphi - (\partial^0 \varphi^\dagger)\varphi) \tag{418}$$

2.9.2 More on Continuity Equation – Complex Field

For $\mathscr{H}(x)$ and $\mathscr{P}(x)$ defined in Eq.(416) and Eq.(417), we are going to prove:

$$\frac{\partial \mathscr{H}}{\partial t} + \nabla \cdot \vec{\mathscr{P}} = 0 \tag{419}$$

$$\begin{aligned}\frac{dH}{dt} &= \frac{d}{dt}\int \mathscr{H}(x) d^3x = \frac{d}{dt}\int \left(\partial^0\varphi^\dagger\partial^0\varphi + \nabla\varphi^\dagger \cdot \nabla\varphi + m^2\varphi^\dagger\varphi\right) d^3x \\ &= \int \left(\ddot{\varphi}^\dagger\dot{\varphi} + \dot{\varphi}^\dagger\ddot{\varphi} + \nabla\dot{\varphi}^\dagger \cdot \nabla\varphi + \nabla\varphi^\dagger \cdot \nabla\dot{\varphi} + m^2\dot{\varphi}^\dagger\varphi + m^2\varphi^\dagger\dot{\varphi}\right) d^3x\end{aligned} \tag{420}$$

We can use the fact that $\varphi(x)$ satisfies the Klein-Gordon equation:

$$\left(\partial^\mu\partial_\mu + m^2\right)\varphi(x) = 0 \to \ddot{\varphi}(x) = \left(\nabla^2 - m^2\right)\varphi(x) \tag{421}$$

$$\left(\partial^\mu\partial_\mu + m^2\right)\varphi^\dagger(x) = 0 \to \ddot{\varphi}^\dagger(x) = \left(\nabla^2 - m^2\right)\varphi^\dagger(x) \tag{422}$$

By plugging Eq.(421) and Eq.(422) into Eq.(420),

$$\begin{aligned}\frac{dH}{dt} &= \int \left(\left(\nabla^2 - m^2\right)\varphi^\dagger(x)\dot{\varphi} + \dot{\varphi}^\dagger\left(\nabla^2 - m^2\right)\varphi(x) + \nabla\dot{\varphi}^\dagger \cdot \nabla\varphi + \nabla\varphi^\dagger \cdot \nabla\dot{\varphi} + m^2\dot{\varphi}^\dagger\varphi + m^2\varphi^\dagger\dot{\varphi}\right) d^3x \\ &= \int \left(\nabla^2\varphi^\dagger\dot{\varphi} + \dot{\varphi}^\dagger\nabla^2\varphi + \nabla\dot{\varphi}^\dagger \cdot \nabla\varphi + \nabla\varphi^\dagger \cdot \nabla\dot{\varphi}\right) d^3x \\ &= \int \nabla \cdot \left(\dot{\varphi}^\dagger\nabla\varphi + \nabla\varphi^\dagger\dot{\varphi}\right) d^3x = -\int \left(\nabla \cdot \vec{\mathscr{P}}\right) d^3x \\ &= -\int_{\text{Boundary}} \vec{\mathscr{P}} \cdot d\vec{a} = 0\end{aligned} \tag{423}$$

From Eq.(423), we see that

$$\begin{aligned}\frac{dH}{dt} &= \int \frac{\partial \mathscr{H}}{\partial t} d^3x = -\int \left(\nabla \cdot \vec{\mathscr{P}}\right) d^3x \\ &\to \int \left(\frac{\partial \mathscr{H}}{\partial t} + \nabla \cdot \vec{\mathscr{P}}\right) d^3x = 0\end{aligned}$$

Thus, we prove

$$\therefore \frac{\partial \mathscr{H}}{\partial t} + \nabla \cdot \vec{\mathscr{P}} = 0 \tag{424}$$

Additionally, from Eq.(423), we see that:

$$\frac{dH}{dt} = 0 \tag{425}$$

as it should be because the Lagrangian does not explicitly depend on time (See Eq.(388)).

2.9.3 Example of Energy-Momentum Tensor, Electromagnetic Field

I found that this section is computationally challenging. But let's go. We consider the Lagrangian of the electromagnetic field:

$$\mathscr{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{426}$$

where

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \qquad \text{and} \qquad A^\mu = \left(\phi, \vec{A}\right) \tag{427}$$

We expand Eq.(426):

$$\begin{aligned}\mathscr{L} &= -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} = -\frac{1}{4} \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) \\ &= -\frac{1}{4} \left(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\nu A^\mu \partial_\mu A_\nu + \partial^\nu A^\mu \partial_\nu A_\mu\right)\end{aligned} \tag{428}$$

Let us calculate the energy-momentum tensor:

$$\begin{aligned}T^{\mu\nu} &= \frac{\partial \mathscr{L}}{\partial(\partial_\mu A_\sigma)} \partial^\nu A_\sigma - g^{\mu\nu}\mathscr{L} \\ &= -\frac{1}{4} \frac{\partial}{\partial(\partial_\mu A_\sigma)} \left(\partial^\alpha A^\beta \partial_\alpha A_\beta - \partial^\alpha A^\beta \partial_\beta A_\alpha - \partial^\beta A^\alpha \partial_\alpha A_\beta + \partial^\beta A^\alpha \partial_\beta A_\alpha\right) \partial^\nu A_\sigma - g^{\mu\nu}\mathscr{L} \\ &= -\frac{1}{4} \left(\frac{\partial(\partial^\alpha A^\beta)}{\partial(\partial_\mu A_\sigma)} \partial_\alpha A_\beta + \partial^\alpha A^\beta \frac{\partial(\partial_\alpha A_\beta)}{\partial(\partial_\mu A_\sigma)} - \frac{\partial(\partial^\alpha A^\beta)}{\partial(\partial_\mu A_\sigma)} \partial_\beta A_\alpha - \partial^\alpha A^\beta \frac{\partial(\partial_\beta A_\alpha)}{\partial(\partial_\mu A_\sigma)} \right. \\ &\left. - \frac{\partial(\partial^\beta A^\alpha)}{\partial(\partial_\mu A_\sigma)} \partial_\alpha A_\beta - \partial^\beta A^\alpha \frac{\partial(\partial_\alpha A_\beta)}{\partial(\partial_\mu A_\sigma)} + \frac{\partial(\partial^\beta A^\alpha)}{\partial(\partial_\mu A_\sigma)} \partial_\beta A_\alpha + \partial^\beta A^\alpha \frac{\partial(\partial_\beta A_\alpha)}{\partial(\partial_\mu A_\sigma)} \right) \partial^\nu A_\sigma \\ &- g^{\mu\nu}\mathscr{L}\end{aligned} \tag{429}$$

To proceed, we need to use the following:

$$\frac{\partial(\partial^\alpha A^\beta)}{\partial(\partial_\mu A_\sigma)} = \frac{\partial(\partial_\gamma A_\rho)}{\partial(\partial_\mu A_\sigma)} g^{\alpha\gamma} g^{\beta\rho} = \delta^\mu_\gamma \delta^\sigma_\rho g^{\alpha\gamma} g^{\beta\rho} = g^{\alpha\mu} g^{\beta\sigma}$$

Then, Eq.(429) becomes:

$$\begin{aligned}T^{\mu\nu} &= -\frac{1}{4} \left(g^{\mu\alpha} g^{\sigma\beta} \partial_\alpha A_\beta + \partial^\alpha A^\beta \delta^\mu_\alpha \delta^\sigma_\beta - g^{\mu\alpha} g^{\sigma\beta} \partial_\beta A_\alpha - \partial^\alpha A^\beta \delta^\mu_\beta \delta^\sigma_\alpha \right. \\ &\left. - g^{\mu\beta} g^{\sigma\alpha} \partial_\alpha A_\beta - \partial^\beta A^\alpha \delta^\mu_\alpha \delta^\sigma_\beta + g^{\mu\beta} g^{\sigma\alpha} \partial_\beta A_\alpha + \partial^\beta A^\alpha \delta^\mu_\beta \delta^\sigma_\alpha \right) \partial^\nu A_\sigma - g^{\mu\nu}\mathscr{L} \\ &= -\frac{1}{4} \left(\partial^\mu A^\sigma + \partial^\mu A^\sigma - \partial^\sigma A^\mu - \partial^\sigma A^\mu - \partial^\sigma A^\mu - \partial^\sigma A^\mu + \partial^\mu A^\sigma + \partial^\mu A^\sigma\right) \partial^\nu A_\sigma \\ &- g^{\mu\nu}\mathscr{L} \\ &= \left(\partial^\sigma A^\mu - \partial^\mu A^\sigma\right) \partial^\nu A_\sigma - g^{\mu\nu}\mathscr{L}\end{aligned} \tag{430}$$

Let us calculate the Hamiltonian density from Eq.(429):

$$\mathscr{H} = T^{00} = \left(\partial^\sigma A^0 - \partial^0 A^\sigma\right)\partial^0 A_\sigma - g^{00}\mathscr{L} \tag{431}$$

Using Eq.(428) in Eq.(431),

$$\begin{aligned}
\mathscr{H} &= \left(\partial^\sigma A^0 - \partial^0 A^\sigma\right)\partial^0 A_\sigma - g^{00}\mathscr{L} \\
&= \left(\partial^\sigma A^0 - \partial^0 A^\sigma\right)\partial^0 A_\sigma + \frac{1}{4}\left(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\nu A^\mu \partial_\mu A_\nu + \partial^\nu A^\mu \partial_\nu A_\mu\right) \\
&= \left(\partial^\sigma A^0 - \partial^0 A^\sigma\right)\partial^0 A_\sigma + \frac{1}{2}\left(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu\right) \\
&= \partial^\sigma A^0 \partial^0 A_\sigma - \partial^0 A^\sigma \partial^0 A_\sigma + \frac{1}{2}\left(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu\right)
\end{aligned} \tag{432}$$

We can expand the summation as follows:

$$\begin{aligned}
\mathscr{H} &= \partial^0 A^0 \partial^0 A_0 + \partial^k A^0 \partial^0 A_k - \partial^0 A^0 \partial^0 A_0 - \partial^0 A^k \partial^0 A_k \\
&\quad + \frac{1}{2}(\partial^0 A^\nu \partial_0 A_\nu + \partial^k A^\nu \partial_k A_\nu) - \frac{1}{2}(\partial^0 A^\nu \partial_\nu A_0 + \partial^k A^\nu \partial_\nu A_k) \\
&= \partial^k A^0 \partial^0 A_k - \partial^0 A^k \partial^0 A_k \\
&\quad + \frac{1}{2}(\partial^0 A^0 \partial_0 A_0 + \partial^0 A^k \partial_0 A_k + \partial^k A^0 \partial_k A_0 + \partial^k A^j \partial_k A_j) \\
&\quad - \frac{1}{2}(\partial^0 A^0 \partial_0 A_0 + \partial^0 A^k \partial_k A_0 + \partial^k A^0 \partial_0 A_k + \partial^k A^j \partial_j A_k)
\end{aligned} \tag{433}$$

Then, Eq.(433) can be simplified to:

$$\begin{aligned}
\mathscr{H} &= \frac{1}{2}(\partial^k A^0 \partial^0 A_k - \partial^0 A^k \partial^0 A_k + \partial^k A^0 \partial_k A_0 - \partial^0 A^k \partial_k A^0) \\
&\quad + \frac{1}{2}(\partial^k A^j \partial_k A_j - \partial^k A^j \partial_j A_k) \\
&= \frac{1}{2}\left(\partial_k A^0 \partial^0 A^k + \sum_{k=1}^{3} \partial^0 A^k \partial^0 A^k - \sum_{k=1}^{3} \partial_k A^0 \partial_k A_0 - \partial^0 A^k \partial_k A^0\right) \\
&\quad + \frac{1}{2}\partial^k A^j (\partial_k A_j - \partial_j A_k)
\end{aligned} \tag{434}$$

Then, Eq.(434) can be written as:

$$\begin{aligned}
\mathscr{H} &= \frac{1}{2}\left(\partial_k A^0 \partial^0 A^k + \sum_{k=1}^{3} \partial^0 A^k \partial^0 A^k + \sum_{k=1}^{3} \partial_k A^0 \partial_k A_0 + \partial^0 A^k \partial_k A^0\right) \\
&\quad + \frac{1}{2}\partial^k A^j (\partial_k A_j - \partial_j A_k) - \left(\sum_{k=1}^{3} \partial_k A^0 \partial_k A_0 + \partial^0 A^k \partial_k A^0\right)
\end{aligned} \tag{435}$$

Eq.(435) can be simplified:

$$
\begin{aligned}
\mathscr{H} &= \frac{1}{2}\left(\nabla\phi \cdot \frac{\partial \vec{A}}{\partial t} + \frac{\partial \vec{A}}{\partial t} \cdot \frac{\partial \vec{A}}{\partial t} + \nabla\phi \cdot \nabla\phi + \frac{\partial \vec{A}}{\partial t} \cdot \nabla\phi\right) \\
&+ \frac{1}{2}\partial^k A^j(\partial_k A_j - \partial_j A_k) - \left(\nabla\phi \cdot \nabla\phi + \frac{\partial \vec{A}}{\partial t} \cdot \nabla\phi\right)
\end{aligned}
\tag{436}
$$

Notice that the second term in Eq.(436) can be written as (For now, let's not use Einstein's summation rule and stick to the conventional summation rule, which is "if two indices are repeated, summation is implied):

$$
\begin{aligned}
\left(\nabla \times \vec{A}\right) \cdot \left(\nabla \times \vec{A}\right) &= (\epsilon_{ijk}\partial_j A_k)\,(\epsilon_{inm}\partial_n A_m) \\
&= \epsilon_{ijk}\epsilon_{inm}\partial_j A_k \partial_n A_m \\
&= (\delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn})\partial_j A_k \partial_n A_m \\
&= \partial_j A_k \partial_j A_k - \partial_j A_k \partial_k A_j \\
&= \partial_j A_k(\partial_j A_k - \partial_k A_j) \\
&= \partial_k A_j(\partial_k A_j - \partial_j A_k)
\end{aligned}
\tag{437}
$$

where the summation indices j and k can be exchanged in the last equality. Recall that

$$
\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \qquad\qquad \vec{B} = \nabla \times \vec{A}
$$

Thus, the Hamiltonian density can be written as:

$$
\mathscr{H} = \frac{1}{2}|\vec{E}|^2 + \frac{1}{2}|\vec{B}|^2 + \vec{E} \cdot \nabla\phi \tag{438}
$$

There is an "unpleasant" term in Eq.(438), which is $\vec{E} \cdot \nabla\phi$. However, this term vanishes once integrated with respect to the volume:

$$
\begin{aligned}
\int \left(\vec{E} \cdot \nabla\phi\right) d^3x &= \int \nabla \cdot \left(\vec{E}\phi\right) d^3x - \int \left(\nabla \cdot \vec{E}\right) \phi d^3x \\
&= \int \left(\vec{E}\phi\right) \cdot d\vec{a} - \int \left(\nabla \cdot \vec{E}\right) \phi d^3x
\end{aligned}
$$

The first term is zero because $\phi \approx 0$ at the boundary, and the second term is zero because of Gauss' law in a vacuum, $\nabla \cdot \vec{E} = 0$. Thus, Eq.(438), when integrated with respect to the volume, is given by:

$$
H = \int \frac{1}{2}\left(|\vec{E}|^2 + |\vec{B}|^2\right) d^3x \tag{439}
$$

which agrees with electrodynamics.

2.9.4 Example of Energy-Momentum Tensor, Dirac Field

Consider the Lagrangian of the Dirac field:

$$\mathscr{L}(x) = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi$$

We have to plug the Lagrangian above into the general expression for the energy-momentum tensor:

$$T^{\mu\nu} = \frac{\partial\mathscr{L}}{\partial(\partial_{\mu}\phi_r)}\partial^{\nu}\phi_r - g^{\mu\nu}\mathscr{L}$$

Then we have the following:

$$\begin{aligned} T^{\mu\nu} &= \frac{\partial\mathscr{L}}{\partial(\partial_{\mu}\psi)}\partial^{\nu}\psi + \frac{\partial\mathscr{L}}{\partial(\partial_{\mu}\bar{\psi})}\partial^{\nu}\bar{\psi} - g^{\mu\nu}\mathscr{L} \\ &= \bar{\psi}i\gamma^{\mu}\partial^{\nu}\psi - g^{\mu\nu}\bar{\psi}\left(i\gamma^{\alpha}\partial_{\alpha} - m\right)\psi \end{aligned} \tag{440}$$

Then the Hamiltonian density is given by:

$$\begin{aligned} \mathscr{H}(x) = T^{00} &= \bar{\psi}i\gamma^{0}\partial^{0}\psi - \bar{\psi}\left(i\gamma^{\alpha}\partial_{\alpha} - m\right)\psi \\ &= \bar{\psi}\left(-i\vec{\gamma}\cdot\nabla + m\right)\psi = \bar{\psi}i\gamma^{0}\partial_{0}\psi = \psi^{\dagger}i\partial_{0}\psi \end{aligned} \tag{441}$$

The momentum density is given by:

$$\mathscr{P}^{i}(x) = T^{0i} = \bar{\psi}i\gamma^{0}\partial^{i}\psi = -\bar{\psi}i\gamma^{0}\partial_{i}\psi = -\psi^{\dagger}i\partial_{i}\psi \tag{442}$$

2.10 Quantization of Fields

Let us consider a system of many particles living in one-dimensional space. Denote x_s as the position operator of an s-th particle. Similarly, we can define p_s as the momentum operator of a s-th particle. Then, Heisenberg's uncertainty principle states:

$$[x_s, p_t] = i\delta_{st} \qquad [x_s, x_t] = 0 \qquad [p_s, p_t] = 0 \tag{443}$$

Now we extend Eq.(443) to fields by taking:

$$x_s \to \phi(x) \qquad p_s \to \pi(x) \tag{444}$$

This means that the position operator is replaced by the position field operator, and the momentum operator is replaced by the momentum field operator. With Eq.(444), Eq.(443) can be written as:

$$[\phi_s(x), \pi_t(y)]_{t_x=t_y} = i\delta_{st}\delta^3(\vec{x}-\vec{y}) \quad [\phi_s(x), \phi_t(y)]_{t_x=t_y} = [\pi_s(x), \pi_t(y)]_{t_x=t_y} = 0 \tag{445}$$

where

$$\pi_t(x) = \frac{\partial\mathscr{L}}{\partial\dot{\phi}_t} \tag{446}$$

2.11 Complex Fields

Suppose we want to solve the complex Klein-Gordon equation:

$$(\partial_\mu \partial^\mu + m^2)\varphi(x) = 0$$

We assume a plane-wave solution:

$$\varphi(x) = q_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} \tag{447}$$

where $q_{\vec{k}}$ is an amplitude as a function of $\vec{k}$. By plugging Eq.(447) into the complex Klein-Gordon equation:

$$\begin{aligned}(\partial_\mu \partial^\mu + m^2) q_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} &= (\partial_0 \partial^0 + \partial_k \partial^k + m^2) q_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} \\ &= (\ddot{q}_{\vec{k}} + (|\vec{k}|^2 + m^2) q_{\vec{k}}) e^{i\vec{k}\cdot\vec{x}} \\ &= (\ddot{q}_{\vec{k}} + \omega^2 q_{\vec{k}}) e^{i\vec{k}\cdot\vec{x}} = 0\end{aligned} \tag{448}$$

where $\omega = \sqrt{|\vec{k}|^2 + m^2}$. Then, we can claim:

$$\begin{aligned}&\ddot{q}_{\vec{k}} + \omega^2 q_{\vec{k}} = 0 \\ &\rightarrow q_{\vec{k}} = a_{\vec{k}} e^{-i\omega t} \qquad \text{or} \qquad q_{\vec{k}} = c_{\vec{k}} e^{i\omega t}\end{aligned} \tag{449}$$

Since the Klein-Gordon equation is a linear equation, its solution is given by a linear combination of Eq.(447):

$$\varphi(x) = \sum_{\vec{k}} q_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t} + c_{\vec{k}} e^{i\omega t} \right) e^{i\vec{k}\cdot\vec{x}} \tag{450}$$

where we put $1/\sqrt{2\omega V}$ for the normalization (V is the volume occupied by the plane wave). Notice that we do not have to sum over the energy ω, because ω is automatically determined once momentum $\vec{k}$ is determined. Keeping in mind the orthogonality relation condition:

$$\frac{1}{V} \int e^{i(\vec{k}-\vec{k'})\cdot\vec{x}} d^3x = \delta_{\vec{k}\vec{k'}} \tag{451}$$

We can prove the following three expressions, which are the Hamiltonian, momentum, and charge, respectively:

$$H = \sum_{\vec{k}} \omega \left(a^\dagger_{\vec{k}} a_{\vec{k}} + c^\dagger_{\vec{k}} c_{\vec{k}} \right) \tag{452}$$

$$\vec{P} = \sum_{\vec{k}} \vec{k} \left(a^\dagger_{\vec{k}} a_{\vec{k}} - c^\dagger_{\vec{k}} c_{\vec{k}} \right) \tag{453}$$

$$Q = \sum_{\vec{k}} q \left(a^\dagger_{\vec{k}} a_{\vec{k}} - c^\dagger_{\vec{k}} c_{\vec{k}} \right) \tag{454}$$

So, let's prove them one by one.

2.11.1 Proof of Eq.(452)

According to Eq.(416), the Hamiltonian density of the complex scalar field is given by:

$$\mathscr{H}(x) = \partial^0\varphi^\dagger\partial^0\varphi + \nabla\varphi^\dagger\cdot\nabla\varphi + m^2\varphi^\dagger\varphi$$

Then the Hamiltonian is given by:

$$H = \int \mathscr{H}(x)d^3x = \int \left(\partial^0\varphi^\dagger\partial^0\varphi + \nabla\varphi^\dagger\cdot\nabla\varphi + m^2\varphi^\dagger\varphi\right)d^3x \tag{455}$$

where, according to Eq.(450),

$$\begin{aligned}\varphi(x) &= \sum_{\vec{k}}\frac{1}{\sqrt{2\omega V}}\left(a_{\vec{k}}e^{-i\omega t} + c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}\\ \varphi^\dagger(x) &= \sum_{\vec{k}}\frac{1}{\sqrt{2\omega V}}\left(a^\dagger_{\vec{k}}e^{i\omega t} + c^\dagger_{\vec{k}}e^{-i\omega t}\right)e^{-i\vec{k}\cdot\vec{x}}\end{aligned} \tag{456}$$

Using Eq.(456), the first term of the integrand in Eq.(455), $\partial^0\varphi^\dagger\partial^0\varphi$ is:

$$\begin{aligned}\partial^0\varphi^\dagger\partial^0\varphi &= \frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(i\omega a^\dagger_{\vec{k'}}e^{i\omega t} - i\omega c^\dagger_{\vec{k'}}e^{-i\omega t}\right)e^{-i\vec{k'}\cdot\vec{x}}\left(-i\omega a_{\vec{k}}e^{-i\omega t} + i\omega c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}\\ &= \frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(\omega^2 a^\dagger_{\vec{k'}}a_{\vec{k}} - \omega^2 a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - \omega^2 c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} + \omega^2 c^\dagger_{\vec{k'}}c_{\vec{k}}\right)e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}\end{aligned} \tag{457}$$

Likewise, using Eq.(456), the second term of the integrand in Eq.(455), $\nabla\varphi^\dagger\cdot\nabla\varphi$ is:

$$\begin{aligned}\nabla\varphi^\dagger\cdot\nabla\varphi &= \frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(a^\dagger_{\vec{k'}}e^{i\omega t} + c^\dagger_{\vec{k'}}e^{-i\omega t}\right)\left(-i\vec{k'}\right)e^{-i\vec{k'}\cdot\vec{x}}\cdot\left(a_{\vec{k}}e^{-i\omega t} + c_{\vec{k}}e^{i\omega t}\right)\left(i\vec{k}\right)e^{i\vec{k}\cdot\vec{x}}\\ &= \frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(a^\dagger_{\vec{k'}}a_{\vec{k}} + a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} + c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} + c^\dagger_{\vec{k'}}c_{\vec{k}}\right)\left(\vec{k'}\cdot\vec{k}\right)e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}\end{aligned} \tag{458}$$

Finally, the third term of the integrand in Eq.(455), $m^2\varphi^\dagger\varphi$ is:

$$\begin{aligned}m^2\varphi^\dagger\varphi &= m^2\frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(a^\dagger_{\vec{k'}}e^{i\omega t} + c^\dagger_{\vec{k'}}e^{-i\omega t}\right)e^{-i\vec{k'}\cdot\vec{x}}\left(a_{\vec{k}}e^{-i\omega t} + c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}\\ &= \frac{m^2}{2\omega V}\sum_{\vec{k},\vec{k}}\left(a^\dagger_{\vec{k'}}a_{\vec{k}} + a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} + c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} + c^\dagger_{\vec{k'}}c_{\vec{k}}\right)e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}\end{aligned} \tag{459}$$

Now we plug the results of Eq.(457), Eq.(458), and Eq.(459) into Eq.(455):

$$H = \frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \int \Big[\omega^2 a^\dagger_{\vec{k'}} a_{\vec{k}} - \omega^2 a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - \omega^2 c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + \omega^2 c^\dagger_{\vec{k'}} c_{\vec{k}}$$
$$\left(a^\dagger_{\vec{k'}} a_{\vec{k}} + a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(\vec{k'} \cdot \vec{k}\right)$$
$$m^2 \left(a^\dagger_{\vec{k'}} a_{\vec{k}} + a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k'}} c_{\vec{k}}\right)\Big] e^{i(\vec{k}-\vec{k'})\cdot\vec{x}} d^3x \tag{460}$$

The x dependence of Eq.(460) is only in $e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}$. Thus, by using Eq.(451), Eq.(460) becomes:

$$H = \frac{1}{2\omega} \sum_{\vec{k},\vec{k'}} \Big[\omega^2 a^\dagger_{\vec{k'}} a_{\vec{k}} - \omega^2 a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - \omega^2 c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + \omega^2 c^\dagger_{\vec{k'}} c_{\vec{k}}$$
$$\left(a^\dagger_{\vec{k'}} a_{\vec{k}} + a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(\vec{k'} \cdot \vec{k}\right)$$
$$m^2 \left(a^\dagger_{\vec{k'}} a_{\vec{k}} + a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k'}} c_{\vec{k}}\right)\Big] \delta_{\vec{k},\vec{k'}}$$

By summing over $\vec{k'}$, the expression above becomes

$$H = \frac{1}{2\omega} \sum_{\vec{k}} \Big[\omega^2 a^\dagger_{\vec{k}} a_{\vec{k}} - \omega^2 a^\dagger_{\vec{k}} c_{\vec{k}} e^{2i\omega t} - \omega^2 c^\dagger_{\vec{k}} a_{\vec{k}} e^{-2i\omega t} + \omega^2 c^\dagger_{\vec{k}} c_{\vec{k}}$$
$$\left(a^\dagger_{\vec{k}} a_{\vec{k}} + a^\dagger_{\vec{k}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k}} c_{\vec{k}}\right) \left(\vec{k} \cdot \vec{k}\right)$$
$$m^2 \left(a^\dagger_{\vec{k}} a_{\vec{k}} + a^\dagger_{\vec{k}} c_{\vec{k}} e^{2i\omega t} + c^\dagger_{\vec{k}} a_{\vec{k}} e^{-2i\omega t} + c^\dagger_{\vec{k}} c_{\vec{k}}\right)\Big]$$

By rearranging the terms:

$$H = \frac{1}{2\omega} \sum_{\vec{k}} \Big[\left(\omega^2 + |\vec{k}|^2 + m^2\right) a^\dagger_{\vec{k}} a_{\vec{k}} + \left(-\omega^2 + |\vec{k}|^2 + m^2\right) a^\dagger_{\vec{k}} c_{\vec{k}} e^{2i\omega t}$$
$$+ \left(-\omega^2 + |\vec{k}|^2 + m^2\right) c^\dagger_{\vec{k}} a_{\vec{k}} e^{-2i\omega t} + \left(\omega^2 + |\vec{k}|^2 + m^2\right) c^\dagger_{\vec{k}} c_{\vec{k}}\Big]$$

However, because $w = \sqrt{|\vec{k}|^2 + m^2}$, the expression above becomes:

$$H = \sum_{\vec{k}} \omega \left[a^\dagger_{\vec{k}} a_{\vec{k}} + c^\dagger_{\vec{k}} c_{\vec{k}}\right] \tag{461}$$

2.11.2 Proof of Eq.(453)

According to Eq.(417), the momentum density of the complex scalar field is given by:

$$\vec{\mathscr{P}}(x) = -(\partial^0 \varphi^\dagger \nabla \varphi + \partial^0 \varphi \nabla \varphi^\dagger)$$

Then, the momentum is given by:

$$\vec{P} = \int \mathscr{P}(x) d^3x = -\int (\partial^0\varphi^\dagger \partial_i \varphi + \partial_i \varphi^\dagger \partial^0 \varphi) d^3x \tag{462}$$

Since we have

$$\varphi(x) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t} + c_{\vec{k}} e^{i\omega t}\right) e^{i\vec{k}\cdot\vec{x}}$$
$$\varphi^\dagger(x) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a^\dagger_{\vec{k}} e^{i\omega t} + c^\dagger_{\vec{k}} e^{-i\omega t}\right) e^{-i\vec{k}\cdot\vec{x}}$$

The first term in the integrand in Eq.(462) is:

$$\begin{aligned}
\partial^0\varphi^\dagger \nabla\varphi &= \frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \left(i\omega a^\dagger_{\vec{k'}} e^{i\omega t} - i\omega c^\dagger_{\vec{k'}} e^{-i\omega t}\right) e^{-i\vec{k'}\cdot\vec{x}} \left(a_{\vec{k}} e^{-i\omega t} + c_{\vec{k}} e^{i\omega t}\right) \left(i\vec{k}\right) e^{i\vec{k}\cdot\vec{x}} \\
&= \frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \left(i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(i\vec{k}\right) e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}
\end{aligned} \tag{463}$$

The second term in the integrand in Eq.(462) is:

$$\begin{aligned}
\nabla\varphi^\dagger \partial^0\varphi &= \frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \left(a^\dagger_{\vec{k'}} e^{i\omega t} + c^\dagger_{\vec{k'}} e^{-i\omega t}\right) \left(-i\vec{k'}\right) e^{-i\vec{k'}\cdot\vec{x}} \left(-i\omega a_{\vec{k}} e^{-i\omega t} + i\omega c_{\vec{k}} e^{i\omega t}\right) e^{i\vec{k}\cdot\vec{x}} \\
&= \frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \left(-i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{-2i\omega t} + i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(-i\vec{k'}\right) e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}
\end{aligned} \tag{464}$$

Now we plug the results of Eq.(463) and Eq.(464) into Eq.(462):

$$\begin{aligned}
\vec{P} = -\frac{1}{2\omega V} \sum_{\vec{k},\vec{k'}} \int \Big[&\left(i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \\
&+ \left(-i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{-2i\omega t} + i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(-i\vec{k'}\right) \Big] e^{i(\vec{k}-\vec{k'})\cdot\vec{x}} d^3x
\end{aligned} \tag{465}$$

The x dependence of Eq.(465) is only in $e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}$. Thus, by using Eq.(451), Eq.(465) becomes:

$$\begin{aligned}
\vec{P} = -\frac{1}{2\omega} \sum_{\vec{k},\vec{k'}} \Big[&\left(i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}} a_{\vec{k}} e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \\
&+ \left(-i\omega a^\dagger_{\vec{k'}} a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{2i\omega t} - i\omega a^\dagger_{\vec{k'}} c_{\vec{k}} e^{-2i\omega t} + i\omega c^\dagger_{\vec{k'}} c_{\vec{k}}\right) \left(-i\vec{k'}\right) \Big] \delta_{\vec{k},\vec{k'}}
\end{aligned}$$

Then, by expanding the expression above, it becomes:

$$\vec{P} = -\frac{1}{2\omega}\sum_{\vec{k},\vec{k}}\Big[\Big(i\omega a^\dagger_{\vec{k}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k}}a_{\vec{k}}e^{-2i\omega t} - i\omega c^\dagger_{\vec{k}}c_{\vec{k}}\Big)\Big(i\vec{k}\Big)$$
$$+\Big(-i\omega a^\dagger_{\vec{k}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k}}c_{\vec{k}}e^{2i\omega t} - i\omega a^\dagger_{\vec{k}}c_{\vec{k}}e^{-2i\omega t} + i\omega c^\dagger_{\vec{k}}c_{\vec{k}}\Big)\Big(-i\vec{k}\Big)\Big]$$

By rearranging the terms, we have

$$\vec{P} = \sum_{\vec{k},\vec{k}}\left(a^\dagger_{\vec{k}}a_{\vec{k}} - c^\dagger_{\vec{k}}c_{\vec{k}}\right)\vec{k} \tag{466}$$

2.11.3 Proof of Eq.(454)

We start with Eq.(397):

$$J^\mu = iq(\varphi^\dagger\partial^\mu\varphi - (\partial^\mu\varphi^\dagger)\varphi)$$

Then, the charge density $\mathscr{Q}$ is given by:

$$\mathscr{Q} = J^0 = iq\left(\varphi^\dagger\partial^0\varphi - (\partial^0\varphi^\dagger)\varphi\right) \tag{467}$$

So, the charge Q is given by:

$$Q = \int \mathscr{Q}d^3x = iq\int\left(\varphi^\dagger\partial^0\varphi - (\partial^0\varphi^\dagger)\varphi\right)d^3x \tag{468}$$

Since we have

$$\varphi(x) = \sum_{\vec{k}}\frac{1}{\sqrt{2\omega V}}\left(a_{\vec{k}}e^{-i\omega t} + c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}$$
$$\varphi^\dagger(x) = \sum_{\vec{k}}\frac{1}{\sqrt{2\omega V}}\left(a^\dagger_{\vec{k}}e^{i\omega t} + c^\dagger_{\vec{k}}e^{-i\omega t}\right)e^{-i\vec{k}\cdot\vec{x}}$$

The first term in the integrand of Eq.(468) is:

$$\varphi^\dagger\partial^0\varphi = \frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(a^\dagger_{\vec{k}'}e^{i\omega t} + c^\dagger_{\vec{k}'}e^{-i\omega t}\right)e^{-i\vec{k}'\cdot\vec{x}}\left(-i\omega a_{\vec{k}}e^{-i\omega t} + i\omega c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}$$

By expanding the expression above,

$$\varphi^\dagger\partial^0\varphi = \frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(-i\omega a^\dagger_{\vec{k}'}a_{\vec{k}} + i\omega a^\dagger_{\vec{k}'}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k}'}a_{\vec{k}}e^{-2i\omega t} + i\omega c^\dagger_{\vec{k}'}c_{\vec{k}}\right)e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} \tag{469}$$

The second term in the integrand of Eq.(468) is:

$$(\partial^0\varphi^\dagger)\varphi = \frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(i\omega a^\dagger_{\vec{k}'}e^{i\omega t} - i\omega c^\dagger_{\vec{k}'}e^{-i\omega t}\right)e^{-i\vec{k}'\cdot\vec{x}}\left(a_{\vec{k}}e^{-i\omega t} + c_{\vec{k}}e^{i\omega t}\right)e^{i\vec{k}\cdot\vec{x}}$$

By expanding the expression above,

$$(\partial^0\varphi^\dagger)\varphi = \frac{1}{2\omega V}\sum_{\vec{k},\vec{k'}}\left(i\omega a^\dagger_{\vec{k'}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}}c_{\vec{k}}\right)e^{i(\vec{k}-\vec{k'})\cdot\vec{x}} \tag{470}$$

Now we plug the results of Eq.(469) and Eq.(470) into Eq.(468):

$$Q = iq\frac{1}{2\omega V}\int\sum_{\vec{k},\vec{k'}}\Big[\left(-i\omega a^\dagger_{\vec{k'}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} + i\omega c^\dagger_{\vec{k'}}c_{\vec{k}}\right) - \left(i\omega a^\dagger_{\vec{k'}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}}c_{\vec{k}}\right)\Big]e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}d^3x \tag{471}$$

The x dependence of Eq.(471) is only in $e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}$. Thus, by using Eq.(451), Eq.(471) becomes:

$$Q = iq\frac{1}{2\omega}\sum_{\vec{k},\vec{k'}}\Big[\left(-i\omega a^\dagger_{\vec{k'}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} + i\omega c^\dagger_{\vec{k'}}c_{\vec{k}}\right) - \left(i\omega a^\dagger_{\vec{k'}}a_{\vec{k}} + i\omega a^\dagger_{\vec{k'}}c_{\vec{k}}e^{2i\omega t} - i\omega c^\dagger_{\vec{k'}}a_{\vec{k}}e^{-2i\omega t} - i\omega c^\dagger_{\vec{k'}}c_{\vec{k}}\right)\Big]\delta_{\vec{k},\vec{k'}}$$

By rearranging the expression above, we have

$$Q = \sum_{\vec{k},\vec{k'}} q\left(a^\dagger_{\vec{k'}}a_{\vec{k}} - c^\dagger_{\vec{k'}}c_{\vec{k}}\right) \tag{472}$$

So, we have proved:

$$H = \sum_{\vec{k}}\omega\left(a^\dagger_{\vec{k}}a_{\vec{k}} + c^\dagger_{\vec{k}}c_{\vec{k}}\right)$$

$$\vec{P} = \sum_{\vec{k}}\vec{k}\left(a^\dagger_{\vec{k}}a_{\vec{k}} - c^\dagger_{\vec{k}}c_{\vec{k}}\right)$$

$$Q = \sum_{\vec{k}} q\left(a^\dagger_{\vec{k}}a_{\vec{k}} - c^\dagger_{\vec{k}}c_{\vec{k}}\right)$$

2.11.4 Interpretation of $a^\dagger_{\vec{k}}$ and $c^\dagger_{\vec{k}}$

Let us interpret $a^\dagger_{\vec{k}}$ and $c^\dagger_{\vec{k}}$ as creation operators and $a_{\vec{k}}$ and $c_{\vec{k}}$ as annihilation operators. Then, we can define the following commutation relations:

$$[a_{\vec{k}}, a^\dagger_{\vec{k'}}] = \delta_{\vec{k},\vec{k'}} \qquad [a_{\vec{k}}, a_{\vec{k'}}] = [a^\dagger_{\vec{k}}, a^\dagger_{\vec{k'}}] = \delta_{\vec{k},\vec{k'}} \tag{473}$$

$$[c_{\vec{k}}, c^\dagger_{\vec{k'}}] = \delta_{\vec{k},\vec{k'}} \qquad [c_{\vec{k}}, c_{\vec{k'}}] = [c^\dagger_{\vec{k}}, c^\dagger_{\vec{k'}}] = \delta_{\vec{k},\vec{k'}} \tag{474}$$

$$[a_{\vec{k}}, c^\dagger_{\vec{k}'}] = [a^\dagger_{\vec{k}}, c_{\vec{k}'}] = [a_{\vec{k}}, c_{\vec{k}'}] = [a^\dagger_{\vec{k}}, c^\dagger_{\vec{k}'}] = 0 \tag{475}$$

Also, just like the creation and annihilation operators were introduced in a quantum harmonic oscillator, we can define the number operator.

$$N = \sum_{\vec{k}} \left(a^\dagger_{\vec{k}} a_{\vec{k}} + c^\dagger_{\vec{k}} c_{\vec{k}} \right) \tag{476}$$

Here $a^\dagger_{\vec{k}} a_{\vec{k}}$ and $c^\dagger_{\vec{k}} c_{\vec{k}}$ are the number operators of particles with momentum $\vec{k}$. Thus, summing over all $\vec{k}$ in Eq.(476) gives the total number of particles in the system.

Let's pause here and think about the meaning of $a^\dagger_{\vec{k}}$ and $c^\dagger_{\vec{k}}$. We defined that $a^\dagger_{\vec{k}}$ is an operator which creates a particle with momentum $\vec{k}$ and $a_{\vec{k}}$ is an operator which annihilates a particle with momentum $\vec{k}$. Then what about $c^\dagger_{\vec{k}}$ and $c_{\vec{k}}$? Let us look at the expression of Q and $\vec{P}$:

$$Q = \sum_{\vec{k}} q \left(a^\dagger_{\vec{k}} a_{\vec{k}} - c^\dagger_{\vec{k}} c_{\vec{k}} \right) = \sum_{\vec{k}} q \left(a^\dagger_{\vec{k}} a_{\vec{k}} \right) + \sum_{\vec{k}} -q \left(c^\dagger_{\vec{k}} c_{\vec{k}} \right)$$

$$\vec{P} = \sum_{\vec{k}} \vec{k} \left(a^\dagger_{\vec{k}} a_{\vec{k}} - c^\dagger_{\vec{k}} c_{\vec{k}} \right) = \sum_{\vec{k}} \vec{k} \left(a^\dagger_{\vec{k}} a_{\vec{k}} \right) + \sum_{\vec{k}} -\vec{k} \left(c^\dagger_{\vec{k}} c_{\vec{k}} \right)$$

Since $+q$ and $+\vec{k}$ is multiplied with $a^\dagger_{\vec{k}} a_{\vec{k}}$, we can think of $a^\dagger_{\vec{k}}$ ($a_{\vec{k}}$) as a creation (annihilation) operator of a *particle* of momentum $\vec{k}$. Likewise, since $-q$ and $-\vec{k}$ is multiplied with $c^\dagger_{\vec{k}} c_{\vec{k}}$, we can think of $c^\dagger_{\vec{k}}$ ($c_{\vec{k}}$) as a creation (annihilation) operator of an *antiparticle* of momentum $-\vec{k}$. Let us investigate further. Consider the Heisenberg equation of motion (See Eq.(62)).

$$\frac{d}{dt} A_H = i[H,\, A_H(t)]$$

Let us take

$$A_H(t) = \varphi(x) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t} + c_{\vec{k}} e^{i\omega t} \right) e^{i\vec{k}\cdot\vec{x}} \tag{477}$$

So we can view that $a_{\vec{k}}$ and $c_{\vec{k}}$ has a time-dependent part, $e^{-i\omega t}$ and $e^{i\omega t}$, respectively. Therefore,

$$\frac{d}{dt} a_{\vec{k}} = -i\omega a_{\vec{k}} = i[H,\, a_{\vec{k}}] \qquad \frac{d}{dt} c_{\vec{k}} = i\omega c_{\vec{k}} = i[H,\, c_{\vec{k}}] \tag{478}$$

Let us introduce an energy eigenket $|E\rangle$ such that

$$H\,|E\rangle = E\,|E\rangle \tag{479}$$

Let's multiply $|E\rangle$ from the right side of Eq.(478):

$$-i\omega a_{\vec{k}} |E\rangle = i[H,\, a_{\vec{k}}] |E\rangle = i(H - E) a_{\vec{k}} |E\rangle \tag{480}$$

$$i\omega c_{\vec{k}} |E\rangle = i[H,\, c_{\vec{k}}] |E\rangle = i(H - E) c_{\vec{k}} |E\rangle \tag{481}$$

Eq.(480) and Eq.(481) can be written as:

$$H a_{\vec{k}} |E\rangle = (E - \omega) a_{\vec{k}} |E\rangle \tag{482}$$

$$H c_{\vec{k}} |E\rangle = (E + \omega) c_{\vec{k}} |E\rangle \tag{483}$$

Eq.(482) tells us that the system loses energy by $\hbar\omega$ if we annihilate a particle with momentum $\vec{k}$. Eq.(483) looks suspicious. It tells us that the system *gains* energy by $\hbar\omega$ if we *annihilate* an antiparticle with momentum $-\vec{k}$. In other words, the system *loses* energy by $-\hbar\omega$ if we *annihilate* an antiparticle with momentum $-\vec{k}$.

Let us look at Eq.(99) again:

$$j^\mu(+e) = +2e|A|^2(E,\, \vec{p}) = -2e|A|^2(-E,\, -\vec{p}) \qquad (E > 0)$$

It tells you the following:

> The negative energy of a particle corresponds to the particle traveling backward in time, and this is the same as a particle with an opposite charge (a.k.a. antiparticle) traveling forward in time.

So we need to have

1. Negative Charge, which we do. (Good)
2. Positive Energy, which we don't. (Bad)
3. Positive Momentum, which we don't. (Bad)

And this is what we have so far:

$$\sum_{\vec{k}} -q \left(c^\dagger_{\vec{k}} c_{\vec{k}} \right) \rightarrow \text{Negative Charge}$$

$$H c_{\vec{k}} |E\rangle = (E + \omega) c_{\vec{k}} |E\rangle \rightarrow \text{Negative Energy}$$

$$\sum_{\vec{k}} -\vec{k} \left(c^\dagger_{\vec{k}} c_{\vec{k}} \right) \rightarrow \text{Negative Momentum}$$

We've already achieved the negative charge. So, we want *positive* energy and *positive* momentum. To do that, we make the following substitution:

$$c_{-\vec{k}} \rightarrow b^\dagger_{\vec{k}} \qquad\qquad c^\dagger_{-\vec{k}} \rightarrow b_{\vec{k}} \tag{484}$$

$$[b_{\vec{k}}, b^\dagger_{\vec{k}'}] = \delta_{\vec{k},\vec{k}'} \qquad\qquad [b_{\vec{k}}, b_{\vec{k}'}] = [b^\dagger_{\vec{k}}, b^\dagger_{\vec{k}'}] = \delta_{\vec{k},\vec{k}'} \tag{485}$$

$$[a_{\vec{k}}, b^\dagger_{\vec{k}'}] = [a^\dagger_{\vec{k}}, b_{\vec{k}'}] = [a_{\vec{k}}, b_{\vec{k}'}] = [a^\dagger_{\vec{k}}, b^\dagger_{\vec{k}'}] = 0 \tag{486}$$

First of all, flipping ($-\vec{k} \to \vec{k}$) deals with the negative momentum issue. Second, the substitution ($c_{-\vec{k}} \to b^\dagger_{\vec{k}}$) means as follows:

> We can restate the annihilation of a particle with $-\hbar\omega$ into creation of a particle with $\hbar\omega$.

Then, with the substitution(475),

$$\begin{aligned}
\varphi(x) &= \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t} + c_{\vec{k}} e^{i\omega t}\right) e^{i\vec{k}\cdot\vec{x}} \\
\to \varphi(x) &= \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t} e^{i\vec{k}\cdot\vec{x}} + c_{\vec{k}} e^{i\omega t} e^{-i\vec{k}\cdot\vec{x}}\right)
\end{aligned}$$

Thus, the new solution is:

$$\varphi(x) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}}\right) \tag{487}$$

Then, we can replace $c_{\vec{k}}$ and $c^\dagger_{\vec{k}}$ as follows:

$$c^\dagger_{-\vec{k}} c_{-\vec{k}} = b_{\vec{k}} b^\dagger_{\vec{k}} = b^\dagger_{\vec{k}} b_{\vec{k}} + 1 \tag{488}$$

Thus, Eq.(452), Eq.(453), and Eq.(454) can be written as:

$$H = \sum_{\vec{k}} \omega \left(a^\dagger_{\vec{k}} a_{\vec{k}} + b^\dagger_{\vec{k}} b_{\vec{k}} + 1\right) \tag{489}$$

$$\vec{P} = \sum_{\vec{k}} \vec{k} \left(a^\dagger_{\vec{k}} a_{\vec{k}} + b^\dagger_{\vec{k}} b_{\vec{k}} + 1\right) \tag{490}$$

$$Q = \sum_{\vec{k}} q \left(a^\dagger_{\vec{k}} a_{\vec{k}} - b^\dagger_{\vec{k}} b_{\vec{k}} - 1\right) \tag{491}$$

Remember to $\vec{k} \to -\vec{k}$ in $\vec{P}$ so that $-\vec{k} c^\dagger_{-\vec{k}} c_{-\vec{k}} \to +\vec{k}\left(b^\dagger_{\vec{k}} b_{\vec{k}} + 1\right)$. In this way, we avoid any possibility that an antiparticle has negative energy (but has an opposite charge).

Of course, the commutation relations Eq.(445) work for $\varphi(x)$ and $\pi(x)$ too:

$$[\varphi(x), \pi(y)]_{t_x = t_y} = i\delta^3(\vec{x} - \vec{y}) \quad [\varphi(x), \varphi(y)]_{t_x = t_y} = [\pi(x), \pi(y)]_{t_x = t_y} = 0 \tag{492}$$

where

$$\pi(y) = \frac{\partial \mathscr{L}}{\partial \dot{\varphi}(y)} = \dot{\varphi}^\dagger(y) \tag{493}$$

2.11.5 Proof of Eq.(492)

$$[\varphi(x),\pi(y)]_{t_x=t_y}=\varphi(x)\pi(y)-\pi(y)\varphi(x) \tag{494}$$

where

$$\varphi(x)=\sum_{\vec{k}}\frac{1}{\sqrt{2\omega V}}\left(a_{\vec{k}}e^{-i\omega t+i\vec{k}\cdot\vec{x}}+b^\dagger_{\vec{k}}e^{i\omega t-i\vec{k}\cdot\vec{x}}\right)$$

$$\pi(x)=\dot{\varphi}^\dagger(y)=\sum_{\vec{k}'}\frac{1}{\sqrt{2\omega V}}\left(i\omega a^\dagger_{\vec{k}'}e^{i\omega t-i\vec{k}'\cdot\vec{y}}-i\omega b_{\vec{k}'}e^{-i\omega t+i\vec{k}'\cdot\vec{y}}\right)$$

Then Eq.(494) can be written as:

$$\begin{aligned}
&[\varphi(x),\pi(y)]_{t_x=t_y}\\
&=\frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(a_{\vec{k}}e^{-i\omega t+i\vec{k}\cdot\vec{x}}+b^\dagger_{\vec{k}}e^{i\omega t-i\vec{k}\cdot\vec{x}}\right)\left(i\omega a^\dagger_{\vec{k}'}e^{i\omega t-i\vec{k}'\cdot\vec{y}}-i\omega b_{\vec{k}'}e^{-i\omega t+i\vec{k}'\cdot\vec{y}}\right)\\
&-\frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(i\omega a^\dagger_{\vec{k}'}e^{i\omega t-i\vec{k}'\cdot\vec{y}}-i\omega b_{\vec{k}'}e^{-i\omega t+i\vec{k}'\cdot\vec{y}}\right)\left(a_{\vec{k}}e^{-i\omega t+i\vec{k}\cdot\vec{x}}+b^\dagger_{\vec{k}}e^{i\omega t-i\vec{k}\cdot\vec{x}}\right)\\
&=\frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(i\omega a_{\vec{k}}a^\dagger_{\vec{k}'}e^{i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}-i\omega a_{\vec{k}}b_{\vec{k}'}e^{-2i\omega t}e^{i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}+i\omega b^\dagger_{\vec{k}}a^\dagger_{\vec{k}'}e^{2i\omega t}e^{-i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}-i\omega b^\dagger_{\vec{k}}b_{\vec{k}'}e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}\right)\\
&-\frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(i\omega a^\dagger_{\vec{k}'}a_{\vec{k}}e^{i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}+i\omega a^\dagger_{\vec{k}'}b^\dagger_{\vec{k}}e^{2i\omega t}e^{-i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}-i\omega b_{\vec{k}'}a_{\vec{k}}e^{-2i\omega t}e^{i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}-i\omega b_{\vec{k}'}b^\dagger_{\vec{k}}e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}\right)\\
&=\frac{1}{2\omega V}\sum_{\vec{k},\vec{k}'}\left(i\omega\left[a_{\vec{k}},a^\dagger_{\vec{k}'}\right]e^{i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}+i\omega\left[b_{\vec{k}},b^\dagger_{\vec{k}'}\right]e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}\right)\\
&=\frac{i}{2V}\sum_{\vec{k},\vec{k}'}\left(\delta_{\vec{k},\vec{k}'}e^{i\vec{k}\cdot\vec{x}-i\vec{k}'\cdot\vec{y}}+\delta_{\vec{k},\vec{k}'}e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}}\right)\\
&=\frac{i}{2V}\sum_{\vec{k}}\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})}+e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}\right)
\end{aligned} \tag{495}$$

Let's pause here and see how to deal with $\sum_{\vec{k}}$. Consider a particle that is confined in a box whose volume is L^3. Let's consider the x-direction. Because the particle cannot exist outside of the box, we know that the particle wave should be "fitted" into the box:

$$n\lambda_x=n\frac{2\pi}{k_x}=L$$

That is,

$$k_x=\frac{2\pi n}{L}\rightarrow\Delta k_x=\frac{2\pi}{L}$$

Thus, we conclude:

$$\Delta k_x\Delta k_y\Delta k_z=\frac{(2\pi)^3}{L^3}$$

So, we get a clue on how to deal with $\sum_{\vec{k}}$:

$$\sum_{\vec{k}} f(\vec{k}) = \frac{1}{\Delta k_x \Delta k_y \Delta k_z} \sum_{\vec{k}} f(\vec{k}) \Delta k_x \Delta k_y \Delta k_z = \frac{L^3}{(2\pi)^3} \int f(\vec{k}) dk_x dk_y dk_z \tag{496}$$

So Eq.(495) becomes:

$$\begin{aligned} [\varphi(x), \pi(y)]_{t_x = t_y} &= \frac{i}{2V} \sum_{\vec{k}} \left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})} + e^{-i\vec{k}\cdot(\vec{x}-\vec{y})} \right) \\ &= \frac{i}{2V} \frac{L^3}{(2\pi)^3} \int \left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})} + e^{-i\vec{k}\cdot(\vec{x}-\vec{y})} \right) dk_x dk_y dk_z \end{aligned} \tag{497}$$

To finish Eq.(497), we use the following identity:

$$\frac{1}{2\pi} \int e^{\pm ik(x-y)} dk = \delta(x-y) \tag{498}$$

V and L^3 cancel each other out. After the integration, Eq.(488) becomes:

$$[\varphi(x), \pi(y)]_{t_x = t_y} = i\delta^{(3)}(\vec{x} - \vec{y}) \tag{499}$$

2.11.6 Explicit Form of $a_{\vec{k}}$ and $b^\dagger_{\vec{k}}$

From Eq.(487), we can explicitly calculate $a_{\vec{k}}$ and $b^\dagger_{\vec{k}}$. Consider the following expression:

$$\begin{aligned} & \frac{e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}} \dot{\varphi}(x) - \frac{i\omega e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}} \varphi(x) \\ &= \frac{e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}} \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(-i\omega a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}} \right) \\ &- \frac{i\omega e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}} \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}} \right) \\ &= \frac{1}{2\omega V} \sum_{\vec{k}} \left(-i\omega a_{\vec{k}} e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{2i\omega t} e^{-i(\vec{k}+\vec{k}')\cdot\vec{x}} \right) \\ &- \frac{1}{2\omega V} \sum_{\vec{k}} \left(i\omega a_{\vec{k}} e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{2i\omega t} e^{-i(\vec{k}+\vec{k}')\cdot\vec{x}} \right) \\ &= \frac{1}{2\omega V} \sum_{\vec{k}} -2i\omega a_{\vec{k}} e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} \\ &= \frac{-i}{V} \sum_{\vec{k}} a_{\vec{k}} e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} \end{aligned} \tag{500}$$

Now we integrate both sides of Eq.(491) with respect to all space:

$$\begin{aligned}\int \left[\frac{e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\dot{\varphi}(x) - \frac{i\omega e^{i\omega t - i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\varphi(x) \right] d^3x &= -i\sum_{\vec{k}} a_{\vec{k}} \frac{1}{V}\int e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} d^3x \\ &= -i\sum_{\vec{k}} a_{\vec{k}}\delta_{\vec{k},\vec{k}'} \\ &= -ia_{\vec{k}'}\end{aligned} \tag{501}$$

By setting

$$\frac{1}{\sqrt{2\omega V}} e^{i\omega t - i\vec{k}'\cdot\vec{x}} = \frac{1}{\sqrt{2\omega V}} e^{ik'\cdot x} = e^*_{k'}(x) \tag{502}$$

We can rewrite Eq.(501) as follows:

$$a_{\vec{k}} = i\int e^*_{k'}(x)\overleftrightarrow{\partial_0}\varphi(x) d^3x \tag{503}$$

where we define:

$$f(x)\overleftrightarrow{\partial_\mu} g(x) = f(x)\frac{\partial g(x)}{\partial x^\mu} - \frac{\partial f(x)}{\partial x^\mu} g(x) \tag{504}$$

We can calculate $b^\dagger_{\vec{k}}$ in a similar way. From Eq.(487), we can explicitly calculate $a_{\vec{k}}$ and $b^\dagger_{\vec{k}}$. Consider the following expression:

$$\begin{aligned}&\frac{e^{-i\omega t + i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\dot{\varphi}(x) + \frac{i\omega e^{-i\omega t + i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\varphi(x) \\ &= \frac{e^{-i\omega t + i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}}\left(-i\omega a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}}\right) \\ &+ \frac{i\omega e^{-i\omega t + i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}}\left(a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}}\right) \\ &= \frac{1}{2\omega V}\sum_{\vec{k}}\left(-i\omega a_{\vec{k}} e^{-2i\omega t} e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}\right) \\ &+ \frac{1}{2\omega V}\sum_{\vec{k}}\left(i\omega a_{\vec{k}} e^{-2i\omega t} e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} + i\omega b^\dagger_{\vec{k}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}\right) \\ &= \frac{1}{2\omega V}\sum_{\vec{k}} 2i\omega b^\dagger_{\vec{k}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}} \\ &= \frac{i}{V}\sum_{\vec{k}} b^\dagger_{\vec{k}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}\end{aligned} \tag{505}$$

Now we integrate both sides of Eq.(505) with respect to all space:

$$\begin{aligned}\int\left[\frac{e^{-i\omega t+i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\dot{\varphi}(x)+\frac{i\omega e^{-i\omega t+i\vec{k}'\cdot\vec{x}}}{\sqrt{2\omega V}}\varphi(x)\right]d^3x &= i\sum_{\vec{k}}b^\dagger_{\vec{k}}\frac{1}{V}\int e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}d^3x\\ &= i\sum_{\vec{k}}b^\dagger_{\vec{k}}\delta_{\vec{k},\vec{k}'}\\ &= ib^\dagger_{\vec{k}'}\end{aligned} \tag{506}$$

By setting

$$\frac{1}{\sqrt{2\omega V}}e^{-i\omega t+i\vec{k}'\cdot\vec{x}}=\frac{1}{\sqrt{2\omega V}}e^{-ik'\cdot x}=e_{k'}(x) \tag{507}$$

We can rewrite Eq.(506) into:

$$b^\dagger_{\vec{k}}=-i\int e_{k'}(x)\overleftrightarrow{\partial_0}\varphi(x)d^3x \tag{508}$$

2.12 Relativistic Normalization

Before jumping into the discussion on the real field, let us discuss how we are going to renormalize the wave function in order to incorporate the special theory of relativity. Suppose there is a particle moving in the x-direction with a constant speed v. Quantum mechanics tells us that integrating the probability density over all space is equal to 1:

$$\int_V \rho d^3x=1 \tag{509}$$

However, we need to ask in which frame the volume element d^3x was calculated in special relativity. Suppose the volume element, d^3x_0, was calculated in the frame which is stationary relative to the moving particle. Also, suppose the volume element, d^3x, was calculated in the frame which is moving in the x-direction relative to the particle. Then, due to Lorentz contraction of the volume element, we have the following:

$$\gamma d^3x_0=d^3x \tag{510}$$

By plugging Eq.(510) into Eq.(509), we see the left-hand side of Eq.(509) is proportional to γ. Thus, the value of the integral (Eq.(509)) depends on choice of a frame. And we don't want that. We want the normalization the same no matter what frame we choose.

Since the left-hand side of Eq.(509) is proportional to γ, we want the right-hand side of Eq.(509) is also proportional to γ. There is a physical observable which is proportional to γ – it is energy:

$$E=\gamma mc^2$$

So we modify Eq.(509) into:

$$\int_V \rho d^3x = 2E \xrightarrow{\text{Rest Frame}} \int_V \rho d^3x_0 = 2mc^2 \tag{511}$$

We can see what Eq.(511) implies that directly calculating ρ for both the Klein-Gordon and Dirac solutions:

Klein-Gordon: $\quad \rho = i\left(\varphi^*\partial_0\varphi - \varphi\partial_0\varphi^*\right) = 2E|N|^2 \quad$ where $\quad \varphi(x) = Ne^{-ip\cdot x}$

Dirac: $\quad \rho = \psi^\dagger\psi = |N|^2 u_r^\dagger(p)u_r(p) = 2E|N|^2 \quad$ where $\quad \psi(x) = Nu_r(p)e^{-ip\cdot x}$

We used Eq.(233) to calculate $u_r^\dagger(p)u_r(p) = 2E$.

So, for both the Klein-Gordon and Dirac solutions, $\rho = 2E|N|^2$. Its spatial integration gives:

$$\int_V \rho d^3x = \int_V 2E|N|^2 d^3x = 2E|N|^2 V = 2E$$

where the last equality comes from Eq.(511). Then we have:

$$N = \frac{1}{\sqrt{V}} \tag{512}$$

which implies there is "one" wave function per box.

2.13 $\left|\vec{k}\right\rangle = (?) \cdot a^\dagger_{\vec{k}} \left|0\right\rangle$

Let us continue the discussion on relativistic normalization. We want to know what we have to multiply in front of $a^\dagger_{\vec{k}}\left|0\right\rangle$ so that we have a relativistically correct normalization. We expect the orthogonality of the kets:

$$\left\langle \vec{l} \middle| \vec{k} \right\rangle \sim \delta^{(3)}(\vec{l} - \vec{k}) \tag{513}$$

Here, $\vec{l}$ and $\vec{k}$ are wave vectors. However, it turns out that $\delta^{(3)}(\vec{l} - \vec{k})$ is not Lorentz-invariant but $E\delta^{(3)}(\vec{l} - \vec{k})$ is! Here is the theorem and its proof.

$E\delta^{(3)}(\vec{l} - \vec{k})$ is Lorentz-invariant.

Proof
Suppose $f(x)$ is a function such that if $f(x) = f(x_0)$ then $x = x_0$.

$$\int g(x)\delta(f(x) - f(x_0))dx = \int g(x)\delta(f(x) - f(x_0))\frac{dx}{df}df = g(x_0)\left(\frac{df}{dx}\right)^{-1}_{x=x_0}$$

But $g(x_0)$ can be written as:

$$g(x_0) = \int g(x)\delta(x - x_0)dx = \left(\frac{df}{dx}\right)_{x=x_0}\int g(x)\delta(f(x) - f(x_0))dx$$

Thus, we conclude that:

$$\delta(x - x_0) = \left(\frac{df}{dx}\right)_{x=x_0} \delta(f(x) - f(x_0)) \tag{514}$$

Now do the following substitution:

$$x = l^1 \quad x_0 = k^1 \quad f(x) = l^{1'} = \beta\gamma\omega + l^1\gamma \quad f(x_0) = k^{1'} = \beta\gamma\omega + k^1\gamma \tag{515}$$

Then Eq.(514) becomes:

$$\delta(l^1 - k^1) = \left(\frac{dl^{1'}}{dl^1}\right)_{l^1=k^1} \delta(l^{1'} - k^{1'}) = \left(\beta\gamma\frac{d\omega}{dl^1} + \gamma\right)_{l^1=k^1} \delta(l^{1'} - k^{1'}) \tag{516}$$

We can use:

$$\frac{d\omega}{dl^1} = \frac{d(\hbar\omega)}{d(\hbar l^1)} = \frac{dE}{dp} = \frac{p}{m} = \frac{\hbar l^1}{\hbar w} = \frac{l^1}{\omega} \tag{517}$$

where we used $E = \hbar\omega$, $p = \hbar l$, and

$$E^2 = p^2 + m^2 \rightarrow 2E\frac{dE}{dp} = 2p \rightarrow \frac{dE}{dp} = \frac{p}{E}$$

Thus, we can write Eq.(516) as:

$$\delta(l^1 - k^1) = \gamma\left(\beta\frac{l^1}{\omega} + 1\right)_{l^1=k^1} \delta(l^{1'} - k^{1'}) = \frac{\gamma}{\omega}\left(\beta k^1 + \omega\right)\delta(l^{1'} - k^{1'}) = \frac{\omega'}{\omega}\delta(l^{1'} - k^{1'}) \tag{518}$$

where we used $\omega' = \beta\gamma k^1 + \gamma\omega$.

Thus, from Eq.(518),

$$\omega\delta(l^1 - k^1) = \omega'\delta(l^{1'} - k^{1'}) \tag{519}$$

So we refine Eq.(513):

$$\left\langle \vec{l} \middle| \vec{k} \right\rangle \sim 2\omega\delta^{(3)}(\vec{l} - \vec{k}) \tag{520}$$

so that the orthogonality condition respects Lorentz invariance. Then we define the ket $\left|\vec{k}\right\rangle$ as:

$$\left|\vec{k}\right\rangle = \sqrt{2\omega}a^\dagger_{\vec{k}}\left|0\right\rangle \tag{521}$$

Then we have:

$$\begin{aligned} \left\langle \vec{l} \middle| \vec{k} \right\rangle &= 2\omega\left\langle 0\right| a_{\vec{l}}a^\dagger_{\vec{k}}\left|0\right\rangle = 2\omega\left\langle 0\right|(a^\dagger_{\vec{k}}a_{\vec{l}} + \delta_{\vec{l},\vec{k}})\left|0\right\rangle \\ &= 2\omega\left\langle 0|0\right\rangle\delta_{\vec{l},\vec{k}} = 2\omega\delta_{\vec{l},\vec{k}} = 2\omega\frac{(2\pi)^3}{V}\delta^{(3)}(\vec{l} - \vec{k}) \end{aligned} \tag{522}$$

If you wonder why the last equality of Eq.(522) holds, here is a brief justification. We start from Eq.(496).

$$\sum_{\vec{k}} f(\vec{k}) = \frac{L^3}{(2\pi)^3}\int f(\vec{k})dk_x dk_y dk_z$$

Then the following holds:

$$f(\vec{l}) = \sum_{\vec{k}} \delta_{\vec{l},\vec{k}} f(\vec{k}) = \frac{L^3}{(2\pi)^3} \int \delta_{\vec{l},\vec{k}} f(\vec{k}) dk_x dk_y dk_z = \int \delta^{(3)}(\vec{l} - \vec{k}) f(\vec{k}) d^3k$$

By comparing the last equality, we conclude:

$$\frac{L^3}{(2\pi)^3} \delta_{\vec{l},\vec{k}} = \delta^{(3)}(\vec{l} - \vec{k}) \tag{523}$$

2.14 Fock Space

Fock space is a mathematical space that allows the number of particles to change. Thus, we are considering the creation or annihilation of particles in the system. Let us calculate the matrix element of a_k in the momentum ket space.

$$\begin{aligned} \langle l| a_{\vec{k}} |m\rangle &= \sqrt{2\omega} \langle l| a_{\vec{k}} a^\dagger_{\vec{m}} |0\rangle = \sqrt{2\omega} \langle l| \left(a^\dagger_{\vec{m}} a_{\vec{k}} + \delta_{\vec{k},\vec{m}} \right) |0\rangle \\ &= \sqrt{2\omega} \delta_{\vec{k},\vec{m}} \langle l|0\rangle = \sqrt{2\omega} \delta_{\vec{k},\vec{m}} \delta_{\vec{l},\vec{0}} \end{aligned} \tag{524}$$

Then the matrix elements of $a^\dagger_k$ in the momentum ket space are given by:

$$\langle l| a^\dagger_{\vec{k}} |m\rangle = \sqrt{2\omega} \delta_{\vec{k},\vec{l}} \delta_{\vec{m},\vec{0}} \tag{525}$$

where we took the Hermitian conjugate of Eq.(524) and exchanged m and l. Using Eq.(524) and Eq.(525), we can calculate the matrix element of $\varphi(x)$:

$$\begin{aligned} \langle l| \varphi(x) |m\rangle &= \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \langle l| \left(a_{\vec{k}} e^{-i\omega t + i\vec{k}\cdot\vec{x}} + b^\dagger_{\vec{k}} e^{i\omega t - i\vec{k}\cdot\vec{x}} \right) |m\rangle \\ &= \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(\langle l| a_{\vec{k}} |m\rangle e^{-ik\cdot x} + \langle l| b^\dagger_{\vec{k}} |m\rangle e^{ik\cdot x} \right) \\ &= \sum_{\vec{k}} \frac{1}{\sqrt{2\omega V}} \left(\sqrt{2\omega} \delta_{\vec{k},\vec{m}} \delta_{\vec{l},\vec{0}} e^{-ik\cdot x} + \sqrt{2\omega} \delta_{\vec{k},\vec{l}} \delta_{\vec{m},\vec{0}} e^{ik\cdot x} \right) \\ &= \frac{1}{\sqrt{V}} \left(\delta_{\vec{l},\vec{0}} e^{-im\cdot x} + \delta_{\vec{m},\vec{0}} e^{il\cdot x} \right) \end{aligned} \tag{526}$$

where the first term corresponds to a particle's plane wave, and the second corresponds to an antiparticle's plane wave. With Eq.(526),

$$\langle 0| \varphi(x) |k\rangle = \frac{1}{\sqrt{V}} e^{-ik\cdot x} \rightarrow \text{Particle's plane wave} \tag{527}$$

$$\langle k| \varphi(x) |0\rangle = \frac{1}{\sqrt{V}} e^{ik\cdot x} \rightarrow \text{Antiparticle's plane wave} \tag{528}$$

If we are dealing with a multi-particle system with n particles and m antiparticles, the state of the system is given by:

$$\begin{aligned}|E\rangle &= |k_1, k_2, ..., k_n; q_1, q_2, ..., q_m\rangle \\ &= \sqrt{2\omega_{k_1}}a^\dagger_{k_1}\sqrt{2\omega_{k_2}}a^\dagger_{k_2}\cdots\sqrt{2\omega_{k_n}}a^\dagger_{k_n}\sqrt{2\omega_{q_1}}b^\dagger_{q_1}\sqrt{2\omega_{q_2}}b^\dagger_{q_2}\cdots\sqrt{2\omega_{q_m}}b^\dagger_{q_m}|0\rangle\end{aligned} \tag{529}$$

where the total energy and the total momentum of the system are given by:

$$\begin{aligned}E &= \omega_{k_1} + \omega_{k_2} + \cdots + \omega_{k_n} + \omega_{q_1} + \omega_{q_2} \cdots + \omega_{q_m} \\ \vec{P} &= \vec{k}_1 + \vec{k}_2 + \cdots + \vec{k}_n + \vec{q}_1 + \vec{q}_2 + \cdots + \vec{q}_n\end{aligned} \tag{530}$$

2.15 Real Field

The Lagrangian of a real field is Eq.(368):

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

Now we use the formula for the energy-momentum tensor (Eq.(413))

$$\begin{aligned}T^{\mu\nu} &= \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_r)}\partial^\nu\phi_r - g^{\mu\nu}\mathscr{L} \\ &= \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi - g^{\mu\nu}\mathscr{L} \\ &= \partial^\mu\phi\partial^\nu\phi - g^{\mu\nu}\frac{1}{2}(\partial_\alpha\phi\partial^\alpha\phi - m^2\phi^2)\end{aligned} \tag{531}$$

Then, the Hamiltonian density is given by:

$$\begin{aligned}\mathscr{H}(x) = T^{00} &= \partial^0\phi\partial^0\phi - g^{00}\frac{1}{2}(\partial_\alpha\phi\partial^\alpha\phi - m^2\phi^2) \\ &= \frac{1}{2}\partial^0\phi\partial^0\phi + \frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{m^2}{2}\phi^2\end{aligned} \tag{532}$$

The momentum density is given by:

$$P^i = T^{0i} = \partial^0\phi\partial^i\phi = -\partial^0\phi\partial_i\phi \tag{533}$$

In fact, the real field ϕ is given by $\varphi = \varphi^\dagger = \phi/\sqrt{2}$. Then, from Eq.(397):

$$J^\mu = iq(\varphi^\dagger\partial^\mu\varphi - (\partial^\mu\varphi^\dagger)\varphi) = \frac{iq}{2}(\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi) = 0$$

since $\phi^\dagger = \phi^* = \phi$. Because the charge current density is zero for a real field, a particle described by a real field is charge-neutral.

2.16 Dirac Field

Eq.(221) and Eq.(228) tell us that a particular solution to the Dirac equation is:

$$\psi(x) = Nu_r(\vec{p})e^{-ip\cdot x} = N^2\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\xi_r\\ \xi_r\end{pmatrix}e^{-ip\cdot x} \rightarrow \text{Particle}$$

$$\psi(x) = Nv_r(\vec{p})e^{ip\cdot x} = N^2\sqrt{m}e^{\frac{1}{2}\vec{\alpha}\cdot\vec{\eta}}\begin{pmatrix}\eta_r\\ -\eta_r\end{pmatrix}e^{ip\cdot x} \rightarrow \text{Antiparticle}$$

Thus, the general solution to the Dirac equation is given by the linear combination of the particular solutions

$$\psi(x) = \sum_{\vec{p},r}\frac{1}{\sqrt{2E_{\vec{p}}V}}\left[a_{\vec{p},r}u_r(\vec{p})e^{-ip\cdot x} + b^\dagger_{\vec{p},r}v_r(\vec{p})e^{ip\cdot x}\right] \tag{534}$$

$$\bar{\psi}(x) = \sum_{\vec{p},r}\frac{1}{\sqrt{2E_{\vec{p}}V}}\left[a^\dagger_{\vec{p},r}\bar{u}_r(\vec{p})e^{ip\cdot x} + b_{\vec{p},r}\bar{v}_r(\vec{p})e^{-ip\cdot x}\right] \tag{535}$$

According to Eq.(441), Eq.(442), and Eq.(55(i)), the energy and momentum of the Dirac field are given by:

$$H = \int \psi^\dagger i\partial_0\psi d^3x \tag{536}$$

$$P^i = \int -\psi^\dagger i\partial_i\psi d^3x \tag{537}$$

$$Q = q\int \psi^\dagger\psi d^3x \tag{538}$$

The analogous version of Eq.(452), Eq.(453), and Eq.(454) is:

$$H = \sum_{\vec{p},r}E_{\vec{p}}\left(a^\dagger_{\vec{p},r}a_{\vec{p},r} - b_{\vec{p},r}b^\dagger_{\vec{p},r}\right) \tag{539}$$

$$\vec{P} = \sum_{\vec{p},r}\vec{p}\left(a^\dagger_{\vec{p},r}a_{\vec{p},r} - b_{\vec{p},r}b^\dagger_{\vec{p},r}\right) \tag{540}$$

$$Q = \sum_{\vec{p},r}q\left(a^\dagger_{\vec{p},r}a_{\vec{p},r} + b_{\vec{p},r}b^\dagger_{\vec{p},r}\right) \tag{541}$$

Again, let us prove them one by one. From Section 1.3, we know that the Dirac equation satisfies the energy momentum relation $E^2 - |\vec{p}|^2 = E^2 - |-\vec{p}|^2 = m^2$. Thus, defining $\vec{p}$ **automatically** specifies E, for a given m. So we can set

$$a_{\vec{p}} = a_p \qquad\qquad E_{\vec{p}} = E_{-\vec{p}} = E_p = E_{-p}$$

where

$a_{\vec{p}}$: An operator that annihilates a particle with momentum $\vec{p}$

a_p: An operator that annihilates a particle with 4-momentum $p^\mu = (E, \vec{p})$

$E_{\vec{p}}$: Energy of a particle with momentum $\vec{p}$

$E_{-\vec{p}}$: Energy of a particle with momentum $-\vec{p}$

E_p: Energy of a particle with 4-momentum $p^\mu = (E, \vec{p})$

E_{-p}: Energy of a particle with 4-momentum $p^\mu = (E, -\vec{p})$

Here, we adopt $a_{\vec{p}}$ and $E_{\vec{p}}$.

2.16.1 Proof of Eq.(539)

We need to calculate the integrand of Eq.(536). To do so, we write $\psi^\dagger$ and $\partial_0\psi$ first.

$$\psi^\dagger = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x}\right] \tag{542}$$

$$\partial_0\psi(x) = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[-iE_{\vec{p}} a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + iE_{\vec{p}} b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x}\right] \tag{543}$$

Then $\psi^\dagger\partial_0\psi$ is given by:

$$\begin{aligned}\psi^\dagger\partial_0\psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x}\right] \\ \times &\left[-iE_{\vec{q}} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} + iE_{\vec{q}} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x}\right]\end{aligned} \tag{544}$$

By expanding Eq.(544),

$$\begin{aligned}\psi^\dagger\partial_0\psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\left[-iE_{\vec{q}} a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x}\right. \\ &+iE_{\vec{q}} a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \\ &-iE_{\vec{q}} b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} \\ &\left.+ iE_{\vec{q}} b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x}\right]\end{aligned} \tag{545}$$

We can use $e^{ip\cdot x} = e^{i(E_{\vec{p}}t-\vec{p}\cdot\vec{x})}$ and $e^{iq\cdot x} = e^{i(E_{\vec{q}}t-\vec{q}\cdot\vec{x})}$. Then, Eq.(545) becomes:

$$\begin{aligned}\psi^\dagger\partial_0\psi = \sum_{\vec{p},\vec{q}}\sum_{r,s}\frac{1}{\sqrt{2E_{\vec{p}}V}}\frac{1}{\sqrt{2E_{\vec{q}}V}}\Big[&-iE_{\vec{q}}a^\dagger_{\vec{p},r}a_{\vec{q},r}u^\dagger_r(\vec{p})u_s(\vec{q})e^{i(E_{\vec{p}}-E_{\vec{q}})t}e^{-i(\vec{p}-\vec{q})\cdot\vec{x}}\\ &+ iE_{\vec{q}}a^\dagger_{\vec{p},r}b^\dagger_{\vec{q},r}u^\dagger_r(\vec{p})v_s(\vec{q})e^{i(E_{\vec{p}}+E_{\vec{q}})t}e^{-i(\vec{p}+\vec{q})\cdot\vec{x}}\\ &-iE_{\vec{q}}b_{\vec{p},r}a_{\vec{q},r}v^\dagger_r(\vec{p})u_s(\vec{q})e^{-i(E_{\vec{p}}+E_{\vec{q}})t}e^{i(\vec{p}+\vec{q})\cdot\vec{x}}\\ &+iE_{\vec{q}}b_{\vec{p},r}b^\dagger_{\vec{q},r}v^\dagger_r(\vec{p})v_s(\vec{q})e^{-i(E_{\vec{p}}-E_{\vec{q}})t}e^{i(\vec{p}-\vec{q})\cdot\vec{x}}\Big]\end{aligned} \tag{546}$$

We take spatial integration to both sides of Eq.(546):

$$\begin{aligned}\int\psi^\dagger\partial_0\psi d^3x = \sum_{\vec{p},\vec{q}}\sum_{r,s}\frac{1}{\sqrt{2E_{\vec{p}}V}}\frac{1}{\sqrt{2E_{\vec{q}}V}}\int\Big[&-iE_{\vec{q}}a^\dagger_{\vec{p},r}a_{\vec{q},r}u^\dagger_r(\vec{p})u_s(\vec{q})e^{i(E_{\vec{p}}-E_{\vec{q}})t}e^{-i(\vec{p}-\vec{q})\cdot\vec{x}}\\ &+ iE_{\vec{q}}a^\dagger_{\vec{p},r}b^\dagger_{\vec{q},r}u^\dagger_r(\vec{p})v_s(\vec{q})e^{i(E_{\vec{p}}+E_{\vec{q}})t}e^{-i(\vec{p}+\vec{q})\cdot\vec{x}}\\ &-iE_{\vec{q}}b_{\vec{p},r}a_{\vec{q},r}v^\dagger_r(\vec{p})u_s(\vec{q})e^{-i(E_{\vec{p}}+E_{\vec{q}})t}e^{i(\vec{p}+\vec{q})\cdot\vec{x}}\\ &+iE_{\vec{q}}b_{\vec{p},r}b^\dagger_{\vec{q},r}v^\dagger_r(\vec{p})v_s(\vec{q})e^{-i(E_{\vec{p}}-E_{\vec{q}})t}e^{i(\vec{p}-\vec{q})\cdot\vec{x}}\Big]d^3x\end{aligned} \tag{547}$$

We can use Eq.(451):

$$\frac{1}{V}\int e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}d^3x = \delta_{\vec{k}\vec{k'}}$$

Then Eq.(547) becomes:

$$\begin{aligned}\int\psi^\dagger\partial_0\psi d^3x = \sum_{\vec{p},\vec{q}}\sum_{r,s}\frac{1}{\sqrt{2E_{\vec{p}}}}\frac{1}{\sqrt{2E_{\vec{q}}}}\Big[&-iE_{\vec{q}}a^\dagger_{\vec{p},r}a_{\vec{q},r}u^\dagger_r(\vec{p})u_s(\vec{q})e^{i(E_{\vec{p}}-E_{\vec{q}})t}\delta_{\vec{p},\vec{q}}\\ &+ iE_{\vec{q}}a^\dagger_{\vec{p},r}b^\dagger_{\vec{q},r}u^\dagger_r(\vec{p})v_s(\vec{q})e^{i(E_{\vec{p}}+E_{\vec{q}})t}\delta_{\vec{p},-\vec{q}}\\ &-iE_{\vec{q}}b_{\vec{p},r}a_{\vec{q},r}v^\dagger_r(\vec{p})u_s(\vec{q})e^{-i(E_{\vec{p}}+E_{\vec{q}})t}\delta_{\vec{p},-\vec{q}}\\ &+iE_{\vec{q}}b_{\vec{p},r}b^\dagger_{\vec{q},r}v^\dagger_r(\vec{p})v_s(\vec{q})e^{-i(E_{\vec{p}}-E_{\vec{q}})t}\delta_{\vec{p},\vec{q}}\Big]\end{aligned} \tag{548}$$

By summing over $\vec{q}$, Eq.(548) becomes:

$$\begin{aligned}\int\psi^\dagger\partial_0\psi d^3x = \sum_{\vec{p}}\sum_{r,s}\frac{1}{2E_{\vec{p}}}\Big[&-iE_{\vec{p}}a^\dagger_{\vec{p},r}a_{\vec{p},s}u^\dagger_r(\vec{p})u_s(\vec{p})\\ &+ iE_{-\vec{p}}a^\dagger_{\vec{p},r}b^\dagger_{-\vec{p},s}u^\dagger_r(\vec{p})v_s(-\vec{p})e^{i(E_{\vec{p}}+E_{-\vec{p}})t}\\ &-iE_{-\vec{p}}b_{\vec{p},r}a_{-\vec{p},s}v^\dagger_r(\vec{p})u_s(-\vec{p})e^{-i(E_{\vec{p}}+E_{-\vec{p}})t}\\ &+iE_{\vec{p}}b_{\vec{p},r}b^\dagger_{\vec{p},s}v^\dagger_r(\vec{p})v_s(\vec{p})\Big]\end{aligned} \tag{549}$$

Note that the second and third terms are zero due to Eq.(238) and Eq.(239). By using Eq.(233) and Eq.(235), Eq.(549) becomes:

$$\begin{aligned}\int \psi^\dagger \partial_0 \psi d^3x &= \sum_{\vec{p}} \sum_{r,s} \frac{1}{2E_{\vec{p}}} \left[-iE_{\vec{p}} a^\dagger_{\vec{p},r} a_{\vec{p},s} 2E_{\vec{p}} \delta_{r,s} + iE_{\vec{p}} b_{\vec{p},r} b^\dagger_{\vec{p},s} 2E_{\vec{p}} \delta_{r,s} \right] \\ &= \sum_{\vec{p},r} -iE_{\vec{p}} \left[a^\dagger_{\vec{p},r} a_{\vec{p},r} - b_{\vec{p},r} b^\dagger_{\vec{p},r} \right]\end{aligned} \tag{550}$$

Thus, the Hamiltonian is given by:

$$H = \int \psi^\dagger i\partial_0 \psi d^3x = \sum_{\vec{p},r} E_{\vec{p}} \left[a^\dagger_{\vec{p},r} a_{\vec{p},r} - b_{\vec{p},r} b^\dagger_{\vec{p},r} \right] \tag{551}$$

2.16.2 Proof of Eq.(540)

We need to calculate the integrand of Eq.(540). To do so, we write $\psi^\dagger$ and $\partial_i \psi$ first.

$$\begin{aligned}\psi^\dagger &= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} \right] \\ \partial_i \psi(x) &= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[-ip_i a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + ip_i b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right]\end{aligned} \tag{552}$$

Then $\psi^\dagger \partial_i \psi$ is given by:

$$\begin{aligned}\psi^\dagger \partial_i \psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} \right] \\ \times &\left[-iq_i a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} + iq_i b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \right]\end{aligned} \tag{553}$$

By expanding Eq.(553),

$$\begin{aligned}\psi^\dagger \partial_i \psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\Big[-iq_i a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} \\ &+ iq_i a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \\ &- iq_i b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} \\ &+ iq_i b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \Big]\end{aligned} \tag{554}$$

We can use $e^{ip\cdot x} = e^{i(E_{\vec{p}}t - \vec{p}\cdot\vec{x})}$ and $e^{iq\cdot x} = e^{i(E_{\vec{q}}t - \vec{q}\cdot\vec{x})}$. Then, Eq.(554) becomes:

$$\begin{aligned}\psi^\dagger \partial_i \psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\Big[-iq_i a^\dagger_{\vec{p},r} a_{\vec{q},r} u^\dagger_r(\vec{p}) u_s(\vec{q}) e^{i(E_{\vec{p}} - E_{\vec{q}})t} e^{-i(\vec{p}-\vec{q})\cdot\vec{x}} \\ &+ iq_i a^\dagger_{\vec{p},r} b^\dagger_{\vec{q},r} u^\dagger_r(\vec{p}) v_s(\vec{q}) e^{i(E_{\vec{p}} + E_{\vec{q}})t} e^{-i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &- iq_i b_{\vec{p},r} a_{\vec{q},r} v^\dagger_r(\vec{p}) u_s(\vec{q}) e^{-i(E_{\vec{p}} + E_{\vec{q}})t} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &+ iq_i b_{\vec{p},r} b^\dagger_{\vec{q},r} v^\dagger_r(\vec{p}) v_s(\vec{q}) e^{-i(E_{\vec{p}} - E_{\vec{q}})t} e^{i(\vec{p}-\vec{q})\cdot\vec{x}} \Big]\end{aligned} \tag{555}$$

We take spatial integration to both sides of Eq.(555):

$$\int \psi^\dagger \partial_i \psi d^3x = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \int \Big[-iq_i a^\dagger_{\vec{p},r} a_{\vec{q},r} u^\dagger_r(\vec{p}) u_s(\vec{q}) e^{i(E_{\vec{p}}-E_{\vec{q}})t} e^{-i(\vec{p}-\vec{q})\cdot\vec{x}} \\ + iq_i a^\dagger_{\vec{p},r} b^\dagger_{\vec{q},r} u^\dagger_r(\vec{p}) v_s(\vec{q}) e^{i(E_{\vec{p}}+E_{\vec{q}})t} e^{-i(\vec{p}+\vec{q})\cdot\vec{x}} \\ - iq_i b_{\vec{p},r} a_{\vec{q},r} v^\dagger_r(\vec{p}) u_s(\vec{q}) e^{-i(E_{\vec{p}}+E_{\vec{q}})t} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} \\ + iq_i b_{\vec{p},r} b^\dagger_{\vec{q},r} v^\dagger_r(\vec{p}) v_s(\vec{q}) e^{-i(E_{\vec{p}}-E_{\vec{q}})t} e^{i(\vec{p}-\vec{q})\cdot\vec{x}} \Big] d^3x \tag{556}$$

We can use Eq.(451):

$$\frac{1}{V}\int e^{i(\vec{k}-\vec{k'})\cdot\vec{x}} d^3x = \delta_{\vec{k}\vec{k'}}$$

Then Eq.(556) becomes:

$$\int \psi^\dagger \partial_i \psi d^3x = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}}} \frac{1}{\sqrt{2E_{\vec{q}}}} \Big[-iq_i a^\dagger_{\vec{p},r} a_{\vec{q},r} u^\dagger_r(\vec{p}) u_s(\vec{q}) e^{i(E_{\vec{p}}-E_{\vec{q}})t} \delta_{\vec{p},\vec{q}} \\ + iq_i a^\dagger_{\vec{p},r} b^\dagger_{\vec{q},r} u^\dagger_r(\vec{p}) v_s(\vec{q}) e^{i(E_{\vec{p}}+E_{\vec{q}})t} \delta_{\vec{p},-\vec{q}} \\ - iq_i b_{\vec{p},r} a_{\vec{q},r} v^\dagger_r(\vec{p}) u_s(\vec{q}) e^{-i(E_{\vec{p}}+E_{\vec{q}})t} \delta_{\vec{p},-\vec{q}} \\ + iq_i b_{\vec{p},r} b^\dagger_{\vec{q},r} v^\dagger_r(\vec{p}) v_s(\vec{q}) e^{-i(E_{\vec{p}}-E_{\vec{q}})t} \delta_{\vec{p},\vec{q}} \Big] \tag{557}$$

By summing over $\vec{q}$, Eq.(557) becomes:

$$\int \psi^\dagger \partial_i \psi d^3x = \sum_{\vec{p}} \sum_{r,s} \frac{1}{2E_{\vec{p}}} \Big[-ip_i a^\dagger_{\vec{p},r} a_{\vec{p},s} u^\dagger_r(\vec{p}) u_s(\vec{p}) \\ - ip_i a^\dagger_{\vec{p},r} b^\dagger_{-\vec{p},s} u^\dagger_r(\vec{p}) v_s(-\vec{p}) e^{i(E_{\vec{p}}+E_{-\vec{p}})t} \\ + ip_i b_{\vec{p},r} a_{-\vec{p},s} v^\dagger_r(p) u_s(-\vec{p}) e^{-i(E_{\vec{p}}+E_{-\vec{p}})t} \\ + ip_i b_{\vec{p},r} b^\dagger_{\vec{p},s} v^\dagger_r(\vec{p}) v_s(\vec{p}) \Big] \tag{558}$$

Note that the second and third terms are zero due to Eq.(238) and Eq.(239). By using Eq.(233) and Eq.(235), Eq.(558) becomes:

$$\begin{aligned} \int \psi^\dagger \partial_i \psi d^3x &= \sum_{\vec{p}} \sum_{r,s} \frac{1}{2E_{\vec{p}}} \left[-ip_i a^\dagger_{\vec{p},r} a_{\vec{p},s} 2E_{\vec{p}} \delta_{r,s} + ip_i b_{\vec{p},r} b^\dagger_{\vec{p},s} 2E_{\vec{p}} \delta_{r,s} \right] \\ &= \sum_{\vec{p},r} -ip_i \left[a^\dagger_{\vec{p},r} a_{\vec{p},r} - b_{\vec{p},r} b^\dagger_{\vec{p},r} \right] \end{aligned} \tag{559}$$

Thus, the momentum is given by:

$$\vec{P} = \int \psi^\dagger i\partial_0 \psi d^3x = \sum_{\vec{p},r} p_i \left[a^\dagger_{\vec{p},r} a_{\vec{p},r} - b_{\vec{p},r} b^\dagger_{\vec{p},r} \right] \tag{560}$$

2.16.3 Proof of Eq.(541)

We need to calculate the integrand of Eq.(541). To do so, we write $\psi^\dagger$ and ψ first.

$$\psi(x) = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right]$$

$$\psi^\dagger(x) = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} \right]$$

Then $\psi^\dagger\psi$ is given by:

$$\begin{aligned}\psi^\dagger\psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\left[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} + b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} \right] \\ &\times \left[a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} + b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \right]\end{aligned} \tag{561}$$

By expanding Eq.(561),

$$\begin{aligned}\psi^\dagger\psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\Big[a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} \\ &+ a^\dagger_{\vec{p},r} u^\dagger_r(\vec{p}) e^{ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \\ &+ b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} a_{\vec{q},r} u_s(\vec{q}) e^{-iq\cdot x} \\ &+ b_{\vec{p},r} v^\dagger_r(\vec{p}) e^{-ip\cdot x} b^\dagger_{\vec{q},r} v_s(\vec{q}) e^{iq\cdot x} \Big]\end{aligned} \tag{562}$$

We can use $e^{ip\cdot x} = e^{i(E_{\vec{p}}t - \vec{p}\cdot\vec{x})}$ and $e^{iq\cdot x} = e^{i(E_{\vec{q}}t - \vec{q}\cdot\vec{x})}$. Then, Eq.(562) becomes:

$$\begin{aligned}\psi^\dagger\psi = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} &\Big[a^\dagger_{\vec{p},r} a_{\vec{q},r} u^\dagger_r(\vec{p}) u_s(\vec{q}) e^{i(E_{\vec{p}}-E_{\vec{q}})t} e^{-i(\vec{p}-\vec{q})\cdot\vec{x}} \\ &+ a^\dagger_{\vec{p},r} b^\dagger_{\vec{q},r} u^\dagger_r(\vec{p}) v_s(\vec{q}) e^{i(E_{\vec{p}}+E_{\vec{q}})t} e^{-i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &+ b_{\vec{p},r} a_{\vec{q},r} v^\dagger_r(\vec{p}) u_s(\vec{q}) e^{-i(E_{\vec{p}}+E_{\vec{q}})t} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &+ b_{\vec{p},r} b^\dagger_{\vec{q},r} v^\dagger_r(\vec{p}) v_s(\vec{q}) e^{-i(E_{\vec{p}}-E_{\vec{q}})t} e^{i(\vec{p}-\vec{q})\cdot\vec{x}} \Big]\end{aligned} \tag{563}$$

We take spatial integration to both sides of Eq.(563):

$$\begin{aligned}\int \psi^\dagger\psi d^3x = \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \int &\Big[a^\dagger_{\vec{p},r} a_{\vec{q},r} u^\dagger_r(\vec{p}) u_s(\vec{q}) e^{i(E_{\vec{p}}-E_{\vec{q}})t} e^{-i(\vec{p}-\vec{q})\cdot\vec{x}} \\ &+ a^\dagger_{\vec{p},r} b^\dagger_{\vec{q},r} u^\dagger_r(\vec{p}) v_s(\vec{q}) e^{i(E_{\vec{p}}+E_{\vec{q}})t} e^{-i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &+ b_{\vec{p},r} a_{\vec{q},r} v^\dagger_r(\vec{p}) u_s(\vec{q}) e^{-i(E_{\vec{p}}+E_{\vec{q}})t} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} \\ &+ b_{\vec{p},r} b^\dagger_{\vec{q},r} v^\dagger_r(\vec{p}) v_s(\vec{q}) e^{-i(E_{\vec{p}}-E_{\vec{q}})t} e^{i(\vec{p}-\vec{q})\cdot\vec{x}} \Big] d^3x\end{aligned} \tag{564}$$

We can use Eq.(451):

$$\frac{1}{V}\int e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}d^3x = \delta_{\vec{k}\vec{k'}}$$

Then Eq.(564) becomes:

$$\begin{aligned}\int \psi^\dagger \psi d^3x = \sum_{\vec{p},\vec{q}}\sum_{r,s}\frac{1}{\sqrt{2E_{\vec{p}}}}\frac{1}{\sqrt{2E_{\vec{q}}}}\Big[&a^\dagger_{\vec{p},r}a_{\vec{q},r}u^\dagger_r(\vec{p})u_s(\vec{q})e^{i(E_{\vec{p}}-E_{\vec{q}})t}\delta_{\vec{p},\vec{q}}\\ &+a^\dagger_{\vec{p},r}b^\dagger_{\vec{q},r}u^\dagger_r(\vec{p})v_s(\vec{q})e^{i(E_{\vec{p}}+E_{\vec{q}})t}\delta_{\vec{p},-\vec{q}}\\ &-b_{\vec{p},r}a_{\vec{q},r}v^\dagger_r(\vec{p})u_s(\vec{q})e^{-i(E_{\vec{p}}+E_{\vec{q}})t}\delta_{\vec{p},-\vec{q}}\\ &+b_{\vec{p},r}b^\dagger_{\vec{q},r}v^\dagger_r(\vec{p})v_s(\vec{q})e^{-i(E_{\vec{p}}-E_{\vec{q}})t}\delta_{\vec{p},\vec{q}}\Big]\end{aligned} \tag{565}$$

By summing over $\vec{q}$, Eq.(565) becomes:

$$\begin{aligned}\int \psi^\dagger \psi d^3x = \sum_{\vec{p}}\sum_{r,s}\frac{1}{2E_{\vec{p}}}\Big[&a^\dagger_{\vec{p},r}a_{\vec{p},s}u^\dagger_r(\vec{p})u_s(\vec{p})\\ &+a^\dagger_{\vec{p},r}b^\dagger_{-\vec{p},s}u^\dagger_r(\vec{p})v_s(-\vec{p})e^{i(E_{\vec{p}}+E_{-\vec{p}})t}\\ &+b_{\vec{p},r}a_{-\vec{p},s}v^\dagger_r(\vec{p})u_s(-\vec{p})e^{-i(E_{\vec{p}}+E_{-\vec{p}})t}\\ &+b_{\vec{p},r}b^\dagger_{\vec{p},s}v^\dagger_r(\vec{p})v_s(\vec{p})\Big]\end{aligned} \tag{566}$$

Note that the second and third terms are zero due to Eq.(238) and Eq.(239). By using Eq.(233) and Eq.(235), Eq.(566) becomes:

$$\begin{aligned}\int \psi^\dagger \psi d^3x &= \sum_{\vec{p}}\sum_{r,s}\frac{1}{2E_{\vec{p}}}\left[a^\dagger_{\vec{p},r}a_{\vec{p},s}2E_{\vec{p}}\delta_{r,s}+b_{\vec{p},r}b^\dagger_{\vec{p},s}2E_{\vec{p}}\delta_{r,s}\right]\\ &= \sum_{\vec{p},r}\left[a^\dagger_{\vec{p},r}a_{\vec{p},r}+b_{\vec{p},r}b^\dagger_{\vec{p},r}\right]\end{aligned} \tag{567}$$

Thus, the charge is given by:

$$Q = q\int \psi^\dagger\psi d^3x = \sum_{\vec{p},r} q\left[a^\dagger_{\vec{p},r}a_{\vec{p},r}+b_{\vec{p},r}b^\dagger_{\vec{p},r}\right] \tag{568}$$

Thus, we have proved:

$$H = \sum_{\vec{p},r}E_{\vec{p}}\left(a^\dagger_{\vec{p},r}a_{\vec{p},r}-b_{\vec{p},r}b^\dagger_{\vec{p},r}\right)$$

$$\vec{P} = \sum_{\vec{p},r}\vec{p}\left(a^\dagger_{\vec{p},r}a_{\vec{p},r}-b_{\vec{p},r}b^\dagger_{\vec{p},r}\right)$$

$$Q = \sum_{\vec{p},r}q\left(a^\dagger_{\vec{p},r}a_{\vec{p},r}+b_{\vec{p},r}b^\dagger_{\vec{p},r}\right)$$

Well, we want a positive sign in front of $b_{\vec{p},r}b^\dagger_{\vec{p},r}$ in the Hamiltonian. Also, we want a negative sign in front of $b_{\vec{p},r}b^\dagger_{\vec{p},r}$ in the charge. These two conditions reflect our previous discussions that an antiparticle has positive energy but an opposite charge. To do so, we impose the following **anticommutation relations**.

$$\begin{aligned}
&\{a_{\vec{p},r}, a^\dagger_{\vec{q},s}\} = \{b_{\vec{p},r}, b^\dagger_{\vec{q},s}\} = \delta_{\vec{p},\vec{q}}\delta_{r,s} \\
&\{a_{\vec{p},r}, a_{\vec{q},s}\} = \{a^\dagger_{\vec{p},r}, a^\dagger_{\vec{q},s}\} = \{b_{\vec{p},r}, b_{\vec{q},s}\} = \{b^\dagger_{\vec{p},r}, b^\dagger_{\vec{q},s}\} = 0 \\
&\{a_{\vec{p},r}, b^\dagger_{\vec{q},s}\} = \{a_{\vec{p},r}, b_{\vec{q},s}\} = \{a^\dagger_{\vec{p},r}, b^\dagger_{\vec{q},s}\} = 0
\end{aligned} \tag{569}$$

Then we have:

$$\{b_{\vec{p},r}, b^\dagger_{\vec{q},s}\} = \delta_{\vec{p},\vec{q}}\delta_{r,s} \rightarrow b_{\vec{p},r}b^\dagger_{\vec{q},s} = -b^\dagger_{\vec{q},s}b_{\vec{p},r} + \delta_{\vec{p},\vec{q}}\delta_{r,s} \tag{570}$$

By plugging Eq.(570) into H, $\vec{P}$, and Q, we have

$$H = \sum_{\vec{p},r} E_{\vec{p}} \left(a^\dagger_{\vec{p},r}a_{\vec{p},r} + b^\dagger_{\vec{p},r}b_{\vec{p},r} - 1 \right) \tag{571}$$

$$\vec{P} = \sum_{\vec{p},r} \vec{p} \left(a^\dagger_{\vec{p},r}a_{\vec{p},r} + b^\dagger_{\vec{p},r}b_{\vec{p},r} - 1 \right) \tag{572}$$

$$Q = \sum_{\vec{p},r} q \left(a^\dagger_{\vec{p},r}a_{\vec{p},r} - b^\dagger_{\vec{p},r}b_{\vec{p},r} + 1 \right) \tag{573}$$

We can derive expressions for $a_{\vec{p},r}$. Here is how. We start with the general expression for $\psi(x)$.

$$\psi(x) = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r}u_r(\vec{p})e^{-ip\cdot x} + b^\dagger_{\vec{p},r}v_r(\vec{p})e^{ip\cdot x} \right]$$

We multiply $\frac{1}{\sqrt{2E_{\vec{q}}V}}e^{iq\cdot x}u^\dagger_s(\vec{q})$ from the right on both sides:

$$\begin{aligned}
&\frac{1}{\sqrt{2E_{\vec{q}}V}}e^{iq\cdot x}u^\dagger_s(\vec{q})\psi(x) \\
&= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}V}} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r}u^\dagger_s(\vec{q})u_r(\vec{p})e^{i(q-p)\cdot x} + b^\dagger_{\vec{p},r}u^\dagger_s(\vec{q})v_r(\vec{p})e^{i(q+p)\cdot x} \right]
\end{aligned} \tag{574}$$

By taking a spatial integration on both sides:

$$\begin{aligned}
&\int \frac{1}{\sqrt{2E_{\vec{q}}V}}e^{iq\cdot x}u^\dagger_s(\vec{q})\psi(x)d^3x \\
&= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}V}} \frac{1}{\sqrt{2E_{\vec{p}}V}} \int \left[a_{\vec{p},r}u^\dagger_s(\vec{q})u_r(\vec{p})e^{i(q-p)\cdot x} + b^\dagger_{\vec{p},r}u^\dagger_s(\vec{q})v_r(\vec{p})e^{i(q+p)\cdot x} \right] d^3x
\end{aligned} \tag{575}$$

We use Eq.(451) on Eq.(575):

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{iq\cdot x} u_s^\dagger(\vec{q})\psi(x) d^3x$$
$$= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{p}}}} \int \left[a_{\vec{p},r} u_s^\dagger(\vec{q}) u_r(\vec{p}) e^{i(E_{\vec{q}}-E_{\vec{p}})t} \delta_{\vec{p},\vec{q}} + b^\dagger_{\vec{p},r} u_s^\dagger(\vec{q}) v_r(\vec{p}) e^{i(E_{\vec{q}}+E_{\vec{p}})t} \delta_{\vec{p},-\vec{q}} \right] d^3x \tag{576}$$

After summing over $\vec{p}$, Eq.(576) becomes:

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{iq\cdot x} u_s^\dagger(\vec{q})\psi(x) d^3x$$
$$= \sum_{r} \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{q}}}} \int \left[a_{\vec{q},r} u_s^\dagger(\vec{q}) u_r(\vec{q}) + b^\dagger_{-\vec{q},r} u_s^\dagger(\vec{q}) v_r(-\vec{q}) e^{i(E_{\vec{q}}+E_{-\vec{q}})t} \delta_{\vec{p},-\vec{q}} \right] d^3x \tag{577}$$

We can use Eq.(233) and Eq.(238) to deal with $u_s^\dagger(\vec{q})u_r(\vec{q})$ and $u_s^\dagger(\vec{q})v_r(-\vec{q})$.

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{iq\cdot x} u_s^\dagger(\vec{q})\psi(x) d^3x = \sum_r \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{q}}}} a_{\vec{q},r} 2E_{\vec{q}} \delta_{r,s} = a_{\vec{q},s}$$

Thus, we have:

$$a_{\vec{p},r} = \int \frac{1}{\sqrt{2E_{\vec{p}}V}} e^{ip\cdot x} u_r^\dagger(\vec{p})\psi(x) d^3x \tag{578}$$

Then $a^\dagger_{\vec{p},r}$ is given by:

$$a^\dagger_{\vec{p},r} = \int \frac{1}{\sqrt{2E_{\vec{p}}V}} e^{-ip\cdot x} \psi^\dagger(x) u_r(\vec{p}) d^3x \tag{579}$$

$b^\dagger_{\vec{p},r}$ is given in a similar manner. Start with $\psi(x)$:

$$\psi(x) = \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right]$$

We multiply $\frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})$ from the right on both sides:

$$\frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x)$$
$$= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}V}} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r} v_s^\dagger(\vec{q}) u_r(\vec{p}) e^{-i(q+p)\cdot x} + b^\dagger_{\vec{p},r} v_s^\dagger(\vec{q}) v_r(\vec{p}) e^{i(p-q)\cdot x} \right] \tag{580}$$

By taking a spatial integration on both sides:

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x) d^3x$$
$$= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}V}} \frac{1}{\sqrt{2E_{\vec{p}}V}} \int \left[a_{\vec{p},r} v_s^\dagger(\vec{q}) u_r(\vec{p}) e^{-i(q+p)\cdot x} + b_{\vec{p},r}^\dagger v_s^\dagger(\vec{q}) v_r(\vec{p}) e^{i(p-q)\cdot x} \right] d^3x \tag{581}$$

We use Eq.(451) on Eq.(581):

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x) d^3x$$
$$= \sum_{\vec{p},r} \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{p}}}} \int \left[a_{\vec{p},r} v_s^\dagger(\vec{q}) u_r(\vec{p}) e^{-i(E_{\vec{q}}+E_{\vec{p}})t} \delta_{\vec{p},-\vec{q}} + b_{\vec{p},r}^\dagger v_s^\dagger(\vec{q}) v_r(\vec{p}) e^{i(E_{\vec{q}}-E_{\vec{p}})t} \delta_{\vec{p},\vec{q}} \right] d^3x \tag{582}$$

After summing over $\vec{p}$, Eq.(582) becomes:

$$\begin{aligned} &\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x) d^3x \\ &= \sum_{r} \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{q}}}} \int \left[a_{-\vec{q},r} v_s^\dagger(\vec{q}) u_r(-\vec{q}) + b_{\vec{q},r}^\dagger v_s^\dagger(\vec{q}) v_r(\vec{q}) \right] d^3x \end{aligned} \tag{583}$$

We can use Eq.(235) and Eq.(239) to deal with $u_s^\dagger(\vec{q})u_r(\vec{q})$ and $u_s^\dagger(\vec{q})v_r(-\vec{q})$.

$$\int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x) d^3x = \sum_{r} \frac{1}{\sqrt{2E_{\vec{q}}}} \frac{1}{\sqrt{2E_{\vec{q}}}} 2E_{\vec{q}} b_{\vec{q},r}^\dagger \delta_{r,s} = b_{\vec{q},s}^\dagger$$

Thus, we have:

$$b_{\vec{q},s}^\dagger = \int \frac{1}{\sqrt{2E_{\vec{q}}V}} e^{-iq\cdot x} v_s^\dagger(\vec{q})\psi(x) d^3x \tag{584}$$

Then $b_{\vec{p},r}$ is given by:

$$b_{\vec{p},r} = \int \frac{1}{\sqrt{2E_{\vec{p}}V}} e^{ip\cdot x} \psi^\dagger(x) v_r(\vec{p}) d^3x \tag{585}$$

2.16.4 Anticommutation Relations between fields

The Dirac fields obey the following anticommutation relations:

$$\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} = \delta^{(3)}(\vec{x}-\vec{y}) \tag{586}$$

$$\{\psi(x), \psi(y)\}_{t_x=t_y} = \{\psi^\dagger(x), \psi^\dagger(y)\}_{t_x=t_y} = 0 \tag{587}$$

2.16.5 Proof of Eq.(586)

From Eq.(534), let us define:

$$\psi(x) = \sum_{\vec{p}} \sum_{r} \frac{1}{\sqrt{2E_{\vec{p}}V}} \left[a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right] \tag{588}$$

$$\psi^\dagger(y) = \sum_{\vec{q}} \sum_{s} \frac{1}{\sqrt{2E_{\vec{q}}V}} \left[a^\dagger_{\vec{q},s} u^\dagger_s(\vec{q}) e^{iq\cdot y} + b_{\vec{q},s} v^\dagger_s(\vec{q}) e^{-iq\cdot y} \right] \tag{589}$$

Then, we calculate the anti-commutator between $\psi_r(x)$ and $\psi_s(y)$:

$$\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \left(a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right) \left(a^\dagger_{\vec{q},s} u^\dagger_s(\vec{q}) e^{iq\cdot y} + b_{\vec{q},s} v^\dagger_s(\vec{q}) e^{-iq\cdot y} \right) \\
&- \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \left(a^\dagger_{\vec{q},s} u^\dagger_s(\vec{q}) e^{iq\cdot y} + b_{\vec{q},s} v^\dagger_s(\vec{q}) e^{-iq\cdot y} \right) \left(a_{\vec{p},r} u_r(\vec{p}) e^{-ip\cdot x} + b^\dagger_{\vec{p},r} v_r(\vec{p}) e^{ip\cdot x} \right)
\end{aligned} \tag{590}$$

By expanding Eq.(590),

$$\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \Big(a_{\vec{p},r} a^\dagger_{\vec{q},s} u_r(\vec{p}) u^\dagger_s(\vec{q}) e^{-i(p\cdot x - q\cdot y)} + a_{\vec{p},r} b_{\vec{q},s} u_r(\vec{p}) v^\dagger_s(\vec{q}) e^{-i(p\cdot x + q\cdot y)} \\
&\qquad + b^\dagger_{\vec{p},r} a^\dagger_{\vec{q},s} v_r(\vec{p}) u^\dagger_s(\vec{q}) e^{i(p\cdot x + q\cdot y)} + b^\dagger_{\vec{p},r} b_{\vec{q},s} v_r(\vec{p}) v^\dagger_s(\vec{q}) e^{i(p\cdot x - q\cdot y)} \Big) \\
&- \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \Big(a^\dagger_{\vec{q},s} a_{\vec{p},r} u^\dagger_s(\vec{q}) u_r(\vec{p}) e^{-i(p\cdot x - q\cdot y)} + a^\dagger_{\vec{q},s} b^\dagger_{\vec{p},r} u^\dagger_s(\vec{q}) v_r(\vec{p}) e^{i(p\cdot x + q\cdot y)} \\
&\qquad + b_{\vec{q},s} a_{\vec{p},r} v^\dagger_s(\vec{q}) u_r(\vec{p}) e^{-i(p\cdot x + q\cdot y)} + b_{\vec{q},s} b^\dagger_{\vec{p},r} v^\dagger_s(\vec{q}) v_r(\vec{p}) e^{i(p\cdot x - q\cdot y)} \Big)
\end{aligned} \tag{591}$$

A common point of confusion in the derivation of the Dirac field anticommutation relations is the apparent contradiction between the matrix form of the first term and the scalar form of the second term in vector notation:

$$\psi\psi^\dagger \sim uu^\dagger \quad \text{(Matrix)}, \quad \text{but} \quad \psi^\dagger\psi \sim u^\dagger u \quad \text{(Scalar?)}$$

This contradiction is resolved by recognizing that the anticommutator $\{\psi_{r,\alpha}(x), \psi^\dagger_{s,\beta}(y)\}$ is calculated for open spinor indices α and β. We do not perform a trace or a contraction over these indices, where $\psi_{r,\alpha}(x)$ represents alpha-th component of $\psi_r(x)$ and $\psi_{s,\beta}(x)$ represents beta-th component of $\psi_s(x)$.

When evaluating the second term $\psi^\dagger_\beta(y)\psi_\alpha(x)$, the coefficient associated with the spinor wavefunction is the product of components:

$$(\psi^\dagger\psi)_{\text{coeff}} \sim v^*_\beta u_\alpha$$

Since the individual components u_α and v_β^* are complex scalars (c-numbers), they commute with each other:

$$v_\beta^* u_\alpha = u_\alpha v_\beta^*$$

We recognize that the term on the right-hand side, $u_\alpha v_\beta^*$, is exactly the definition of the matrix element (α, β) of the outer product $uv^\dagger$:

$$u_\alpha v_\beta^* \equiv (uv^\dagger)_{\alpha\beta}$$

Here is more "visual" way of understanding it:

$$u_r v_s^\dagger = v_s^\dagger u_r \tag{592}$$

In linear algebra, Eq.(592) does not make any sense, because the LHS is a matrix and the RHS is a number. However, we have to emphasize that u_r and $v_s^\dagger$ are from different vector spaces – u_r describes a particle and $v_s^\dagger$ describes an antiparticle. Let

$$u_r = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \qquad\qquad v_s = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} \tag{593}$$

By plugging Eq.(593) into Eq.(592), the LHS becomes:

$$\begin{aligned} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} &= \left(\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} v_1^* \quad \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} v_2^* \quad \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} v_3^* \quad \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} v_4^* \right) \\ &= \begin{pmatrix} u_1v_1^* & u_1v_2^* & u_1v_3^* & u_1v_4^* \\ u_2v_1^* & u_2v_2^* & u_2v_3^* & u_2v_4^* \\ u_3v_1^* & u_3v_2^* & u_3v_3^* & u_3v_4^* \\ u_4v_1^* & u_4v_2^* & u_4v_3^* & u_4v_4^* \end{pmatrix} \end{aligned} \tag{594}$$

The RHS becomes:

$$\begin{aligned} \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} &= \begin{pmatrix} \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} u_1 \\ \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} u_2 \\ \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} u_3 \\ \begin{pmatrix} v_1^* & v_2^* & v_3^* & v_4^* \end{pmatrix} u_4 \end{pmatrix} \\ &= \begin{pmatrix} u_1v_1^* & u_1v_2^* & u_1v_3^* & u_1v_4^* \\ u_2v_1^* & u_2v_2^* & u_2v_3^* & u_2v_4^* \\ u_3v_1^* & u_3v_2^* & u_3v_3^* & u_3v_4^* \\ u_4v_1^* & u_4v_2^* & u_4v_3^* & u_4v_4^* \end{pmatrix} \end{aligned} \tag{595}$$

You see that Eq(594) is equivalent to Eq.(595). Thus, we accept Eq.(592). Thus, we can proceed simplifying Eq.(591):

$$
\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \Big(\{a_{\vec{p},r}, a^\dagger_{\vec{q},s}\} u_r(\vec{p}) u_s^\dagger(\vec{q}) e^{-i(p\cdot x - q\cdot y)} + \{a_{\vec{p},r}, b_{\vec{q},s}\} u_r(\vec{p}) v_s^\dagger(\vec{q}) e^{-i(p\cdot x + q\cdot y)} \\
&\qquad + \{b^\dagger_{\vec{p},r}, a^\dagger_{\vec{q},s}\} v_r(\vec{p}) u_s^\dagger(\vec{q}) e^{i(p\cdot x + q\cdot y)} + \{b^\dagger_{\vec{p},r}, b_{\vec{q},s}\} v_r(\vec{p}) v_s^\dagger(\vec{q}) e^{i(p\cdot x - q\cdot y)} \Big)
\end{aligned}
\tag{596}
$$

Due to Eq.(569), the second and the third term are zero. Thus, Eq.(596) becomes:

$$
\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \Big(\{a_{\vec{p},r}, a^\dagger_{\vec{q},s}\} u_r(\vec{p}) u_s^\dagger(\vec{q}) e^{-i(p\cdot x - q\cdot y)} + \{b^\dagger_{\vec{p},r}, b_{\vec{q},s}\} v_r(\vec{p}) v_s^\dagger(\vec{q}) e^{i(p\cdot x - q\cdot y)} \Big) \\
&= \sum_{\vec{p},\vec{q}} \sum_{r,s} \frac{1}{\sqrt{2E_{\vec{p}}V}} \frac{1}{\sqrt{2E_{\vec{q}}V}} \Big(\delta_{\vec{p},\vec{q}} \delta_{r,s} u_r(\vec{p}) u_s^\dagger(\vec{q}) e^{-i(p\cdot x - q\cdot y)} + \delta_{\vec{p},\vec{q}} \delta_{r,s} v_r(\vec{p}) v_s^\dagger(\vec{q}) e^{i(p\cdot x - q\cdot y)} \Big)
\end{aligned}
\tag{597}
$$

By summing over $\vec{q}$ and s,

$$
\begin{aligned}
&\{\psi_r(x), \psi_s^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p}} \frac{1}{2E_{\vec{p}}V} \left(\sum_r u_r(\vec{p}) u_r^\dagger(\vec{p}) e^{-ip\cdot(x-y)} + \sum_r v_r(\vec{p}) v_r^\dagger(\vec{p}) e^{ip\cdot(x-y)} \right)
\end{aligned}
\tag{598}
$$

We should impose $t_x = t_y$ and can use $\gamma^0\gamma^0 = I$:

$$
\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p}} \frac{1}{2E_{\vec{p}}V} \left(\sum_r u_r(\vec{p}) u_r^\dagger(\vec{p}) \gamma^0\gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} + \sum_r v_r(\vec{p}) v_r^\dagger(\vec{p}) \gamma^0\gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \right) \\
&= \sum_{\vec{p}} \frac{1}{2E_{\vec{p}}V} \left(\sum_r u_r(\vec{p}) \bar{u}_r(\vec{p}) \gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} + \sum_r v_r(\vec{p}) \bar{v}_r(\vec{p}) \gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \right)
\end{aligned}
\tag{599}
$$

We note that

$$
\sum_{\vec{p}} v_r(\vec{p}) \bar{v}_r(\vec{p}) \gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \sum_{\vec{p}} v_r(-\vec{p}) \bar{v}_r(-\vec{p}) \gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})}
$$

Thus, Eq.(599) can be simplified into:

$$
\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p}} \frac{1}{2E_{\vec{p}}V} \left(\sum_r u_r(\vec{p}) \bar{u}_r(\vec{p}) + \sum_r v_r(\vec{p}) \bar{v}_r(\vec{p}) \right) \gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})}
\end{aligned}
\tag{600}
$$

We can use Eq.(248) and Eq.(254) to simplify further:

$$
\begin{aligned}
&\{\psi(x), \psi^\dagger(y)\}_{t_x=t_y} \\
&= \sum_{\vec{p}} \frac{1}{2E_{\vec{p}}V} \left(\gamma^0 p_0 + \gamma^k p_k + \gamma^0 p_0 - \gamma^k p_k\right) \gamma^0 e^{-ip\cdot(x-y)} \\
&= \sum_{\vec{p}} \frac{e^{-ip\cdot(x-y)}}{V} = \frac{1}{\Delta p_x \Delta p_y \Delta p_z} \sum_{\vec{p}} \frac{e^{-ip\cdot(x-y)}}{V} \Delta p_x \Delta p_y \Delta p_z \\
&= \frac{V}{(2\pi)^3} \int \frac{e^{-ip\cdot(x-y)}}{V} d^3p = \delta^{(3)}(\vec{x} - \vec{y})
\end{aligned}
\tag{601}
$$