Lecture Notes

PHY 853 - Advanced Quantum Mechanics

http://mo.pa.msu.edu/phy853

Instructor: Scott Pratt, prattsc@pa.msu.edu

These notes are **NOT** meant to be a substitute for the book. The course will follow the order of material in the text. These notes are only meant to outline what was covered in class and to provide something students can print and annotate rather than taking full lecture notes.

Contents

1 Review of Classical Relativity

1.1 γ Factors and Such

The first part of the course will cover the Dirac equation, and later the material will shift to the basics and foundations of relativistic field theory. All such material will seem intimidating without a firm grasp of the principles and formalism of the special theory of relativity in the context of classical physics, especially the use of projections. First, we will review the standard arguments for Lorentz length contraction and for time dilation, i.e., we will demonstrate how a meter stick moving with velocity $\boldsymbol{v}$ appears shorter by a factor $\mathbf{1}/\boldsymbol{\gamma}$, where $\boldsymbol{\gamma = 1/\sqrt{1-(v/c)^2}}$, and will also discuss how a moving clock has ticks separated by extended times, $\boldsymbol{\gamma}$.

Both these results stem from the basic postulate of special relativity, that the speed of light is the same in all reference frames. First, we consider time dilation. Consider two mirrors separated by distance $\boldsymbol{L_0}$. To an observer in the frame of the mirror, the time for light to bounce back and forth from the mirrors is $\boldsymbol{t_0 = 2L_0/c}$. Now let the mirror move perpendicular to the path of the light with speed $\boldsymbol{v}$. An observer in the laboratory sees the same light pulse move with speed $\boldsymbol{c}$, but that the path is longer than $\boldsymbol{2L_0}$ because the return point will have moved a distance $\boldsymbol{vt_{\rm lab}}$, making the entire distance equal to

$$\boldsymbol{L = 2\sqrt{L_0^2 + (vt_{\rm lab}/2)^2}.} \tag{1.1}$$

The time for the return in the lab frame is thus

$$\boldsymbol{t_{\rm lab} = \frac{2\sqrt{L_0^2 + (vt_{\rm lab}/2)^2}}{c}.} \tag{1.2}$$

Replacing $\boldsymbol{L_0}$ with $\boldsymbol{ct_0/2}$, one can solve for $\boldsymbol{t_{\rm lab}}$ in terms of $\boldsymbol{t_0}$,

$$\boldsymbol{t_{\rm lab} = \gamma t_0, \quad \gamma \equiv \frac{1}{\sqrt{1-(v/c)^2}}.} \tag{1.3}$$

Thus, all moving clocks run slow. The most basic manifestation of this considers the lifetime of radioactive decays, which are extended by the factor $\boldsymbol{\gamma}$. Note that for the future, we will often work in units where $\boldsymbol{c = 1}$ to save ink and eyestrain.

Secondly, we consider the same set of mirrors, but instead let the apparatus move parallel to the path of the light (perpendicular to the plane of the mirrors). Again, the observer moving with the stick sees the time between pulses as $\boldsymbol{2L_0/c}$. The laboratory observer measures,

$$\boldsymbol{t_{\rm lab} = \frac{L_{\rm lab}}{c+v} + \frac{L_{\rm lab}}{c-v} = \frac{2cL_{\rm lab}}{c^2-v^2} = \frac{2\gamma^2 L_{\rm lab}}{c},} \tag{1.4}$$

where here we have allowed the length of the meter stick to change from the value observed in the frame of the stick. Using the fact that $\boldsymbol{t_{\rm lab} = \gamma t_0 = 2\gamma L_0/c}$,

$$\begin{aligned} \boldsymbol{2\gamma\frac{L_0}{c}} &= \boldsymbol{2\gamma^2\frac{L_{\rm lab}}{c},} \\ \boldsymbol{L_{\rm lab}} &= \boldsymbol{\frac{L_0}{\gamma}.} \end{aligned} \tag{1.5}$$

Thus moving meter sticks appear shorter.
One has to be careful to note that the two simple expressions apply for very specific circumstances. The expression for time dilation, $t_{\rm lab} = \gamma t_0$, only applies when the time separates two events which occur at the same location in the frame of the moving observer measuring t_0. Also, the expression for length contraction, $L_{\rm lab} = L_0/\gamma$, is applicable only when the distance between the moving ends are measured simultaneously in time according to the observer measuring L_0. These subtleties are illustrated by the ladder paradox. Imagine a runner moving at 60% of the speed of light carrying a 10 foot ladder and moving through a 10 foot garage that has doors at both ends. The gamma factor is 1.25. Thus, an observer in the frame of the garage sees an 8-foot ladder moving through the garage, and could in principle close the garage at both ends trapping the ladder completely inside. The runner sees an 8-foot garage, and believes there was a moment when the front of the ladder had gone completely through the garage while the back of the ladder had not penetrated the front door. The paradoxical question is "Did the ladder fit?". The answer has to do with the simultaneity of two events. In this case it could be lights flashing at the front and back of the ladder. If the runner thinks the lights flash simultaneously, indeed both lights flash outside the garage and on opposite sides. The observer in the garage frame agrees with the assessment that the lights flashed outside the garage, but instead thinks that the light at the back of the ladder blinked first and that the light at the front of the ladder blinked later. Conversely, if the lights blinked such that the observer in the garage frame thought they were simultaneous, the two lights could have both flashed inside the garage. However, the runner would have recorded the light at the front of the ladder blinking first.
The ladder paradox underscores the importance of thinking of times and distances as describing the difference between two events, i.e., times and displacements are always relative to something. There are other famous paradoxes in relativity that also lead to a better understanding of the essence of the theory, such as the twin paradox or whether the radius of a rotating wheel contracts. However, the latter two involve acceleration and are thus related to the general theory of relativity, which is not considered here, but would be considered in a course on gravity.

1.2 Lorentz Transformations

Space and time are even mixed together in non-relativistic (Newtonian) transformations. For instance, consider an event that occurs at time t and position x in the laboratory frame (Here we consider only one spatial dimension). In a moving frame, the event occurs at:

$$x_{\rm lab} = x_0 + vt_0, \quad t_{\rm lab} = t_0, \tag{1.6}$$

in a Newtonian transformation. For a relativistic transformation, we assume a more general linear form,

$$x_{\rm lab} = Ax_0 + Bvt_0, \tag{1.7}$$
$$t_{\rm lab} = Ct_0 + Dvx_0, \tag{1.8}$$

where A, B, C and D are functions of v^2. The powers of v are required by parity considerations. The inverse transformation should look the same, but with $v \to -v$,

$$x_0 = Ax_{\rm lab} - Bvt_{\rm lab}, \tag{1.9}$$
$$t_0 = Ct_{\rm lab} - Dvx_{\rm lab}.$$

First, we consider the decay of a relativistic particle which passes by the point ($x = t = 0$) in both frames. Since the particle does not move in the particle frame ($x_0 = 0$). The transformations becomes

$$\begin{aligned} t_{\rm lab} &= Ct_0, \\ x_{\rm lab} &= Bvt_0. \end{aligned} \tag{1.10}$$

The expression for time dilation in the previous section then yields,

$$C = \gamma. \tag{1.11}$$

The decay occurs at the position $x_{\rm lab} = vt_{\rm lab}$, and since $t_{\rm lab} = \gamma t_0$, one finds

$$B = \gamma. \tag{1.12}$$

To solve for the other two coefficients, consider the ladder paradox. Assume a light at the back end of the latter blinks at a time $x = t = 0$, and at the front end of the ladder a light blinks simultaneously (in the frame of the ladder) at the space-time point, $x_0 = L_0, t_0 = 0$. In the laboratory frame the lights blink at times different by an amount DvL_0. The position at which the light blinks is

$$x_{\rm lab} = L_{\rm lab} + vt_{\rm lab}, \tag{1.13}$$

where $L_{\rm lab}$ is the length of the ladder. Since the apparent length of the ladder is shrunk by a factor γ, $L_{.\rm lab} = x_0/\gamma$, and

$$x_{\rm lab} = x_0/\gamma + vt_{\rm lab}, \tag{1.14}$$

Rearranged,

$$x_0 = \gamma x_{\rm lab} - \gamma v t_{\rm lab}. \tag{1.15}$$

This gives $A = \gamma$. To solve for D take the expressions, Eq.s (1.7,1.9,1.10) in the Lorentz transformations, and solve for D. One finds $D = \gamma$.
The expressions for the coefficients defined in Eq.s (1.7-1.10)) can be summed up as a matrix equation,

$$\begin{aligned} x^\alpha &= L^\alpha_{\ \beta} x'^\beta, \\ L &= \begin{pmatrix} \gamma & \gamma v \\ \gamma v & \gamma \end{pmatrix}, \end{aligned} \tag{1.16}$$

Here, the indices α are either 0 or 1, with "0" referring to time component and "1" referring to the 'x" component. If we included y and z, there would be four components representing the coordinate of an event, t, x, y, z. These four components make up a "four-vector". The Lorentz matrix L performs a boost along the x axis and leaves the other two dimensions unchanged. Thus, for a boost along the x-axis, the 4×4 matrix becomes

$$L = \begin{pmatrix} \gamma & \gamma v & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.17}$$

For rotations of a coordinate system, all three vectors are transformed according to the same rotation matrix $\boldsymbol{U}$, $\boldsymbol{r'_i = U_{ij} r_j}$, regardless of whether the vector $\vec{\boldsymbol{r}}$ refers to a spatial coordinate or a velocity or momentum. Similarly, for boosts all four-vectors are transformed the same,

$$\boldsymbol{r^\alpha \;\; = \;\; L^\alpha_\beta r^\beta}, \tag{1.18}$$

regardless of whether the quantity $\boldsymbol{r}$ represents the space-time coordinate of an event or a momenta (in which case the zero-th component is the energy). For our notes we will stick to the convention that greek indices label the components of four-vectors, while roman indices denote the components of three-vectors, i.e., $\boldsymbol{\alpha = 0, 1, 2, 3}$ and $\boldsymbol{i = 1, 2, 3}$.. The choice of making the indices upper vs. lower will be explained later.
Due to the identity, $\boldsymbol{\gamma^2 - \gamma^2 v^2 = 1}$, one can express γ and $\boldsymbol{\gamma v}$ as $\boldsymbol{\cosh \eta}$ and $\boldsymbol{\sinh \eta}$ respectively. The Lorentz matrix then becomes

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{\cosh \eta} & \boldsymbol{\sinh \eta} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{\sinh \eta} & \boldsymbol{\cosh \eta} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{pmatrix}, \tag{1.19}$$

which illustrates the similarity of the Lorentz transformations to rotations by an imaginary angle. Whereas for a rotation, one finds that $\boldsymbol{x_1 y_1 + x_2 y_2 + x_3 y_3 = \vec{x} \cdot \vec{y}}$ is invariant to rotations, Lorentz transformations leave the quantity,

$$\boldsymbol{x \cdot y \equiv x_0 y_0 - x_1 y_1 - x_2 y_2 - x_3 y_3} \tag{1.20}$$

unchanged. The "dot-product" of four vectors is known as a Lorentz invariant. The dot-product is also invariant to rotations in $\boldsymbol{3}-$space, and the combinations of rotations and boosts is known as the *Lorentz group*. For a group, any combinations of transformations can be expressed as a single transformation. If one adds parity and time-reversal, it is known as the full Lorentz group, and if one adds translational symmetry, it becomes the Poincaré group. Rotations and boosts are not in separable groups, i.e., if one performs two boosts, the result is a boost plus a rotation. If the boosts had formed a group by themselves, any two boosts would have been equivalent to a single boost.

EXAMPLE:
Consider two successive boosts along the $\boldsymbol{x}$ direction. The first defined by $\boldsymbol{\gamma v = \sinh \eta_1}$, and the second one with $\boldsymbol{\gamma v = \sinh \eta_2}$. Show that the combination is equivalent to one boost with $\boldsymbol{\gamma v = \sinh(\eta_1 + \eta_2)}$.

Writing down the product of the two Lorentz matrices (2-dimensions is sufficient),

$$\begin{aligned} &\begin{pmatrix} \boldsymbol{\cosh \eta_2} & \boldsymbol{\sinh \eta_2} \\ \boldsymbol{\sinh \eta_2} & \boldsymbol{\cosh \eta_2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\cosh \eta_1} & \boldsymbol{\sinh \eta_1} \\ \boldsymbol{\sinh \eta_1} & \boldsymbol{\cosh \eta_1} \end{pmatrix} \\ &\qquad = \begin{pmatrix} \boldsymbol{\cosh \eta_2 \cosh \eta_1 + \sinh \eta_2 \sinh \eta_1} & \boldsymbol{\cosh \eta_2 \sinh \eta_1 + \sinh \eta_2 \cosh \eta_1} \\ \boldsymbol{\cosh \eta_2 \sinh \eta_1 + \sinh \eta_2 \cosh \eta_1} & \boldsymbol{\cosh \eta_2 \cosh \eta_1 + \sinh \eta_2 \sinh \eta_1} \end{pmatrix} \\ &\qquad = \begin{pmatrix} \boldsymbol{\cosh(\eta_1 + \eta_2)} & \boldsymbol{\sinh(\eta_1 + \eta_2)} \\ \boldsymbol{\sinh(\eta_1 + \eta_2)} & \boldsymbol{\cosh(\eta_1 + \eta_2)} \end{pmatrix}, \end{aligned}$$

where double-angle formulas were applied for the last step. For high-energy phenomenology the the quantities η are referred to as "rapidities" when the boosts are along the beam axis. The simplicity of rapidities comes from the fact that they add like Newtonian velocities. However, this simple addition only works when the rapidities are defined along a single axis, i.e., if one defines $\sinh \eta_x = \gamma v_x, \cdots$, for all three dimensions, the addition formulas break down due to the non-commutation of boosts along different directions.

1.3 Invariants and the metric tensor $g^{\alpha\beta}$

As shown previously, the dot product of two vectors,

$$A^\alpha g_{\alpha\beta} B^\beta, \tag{1.21}$$

$$g_{\alpha\beta} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.22}$$

is invariant to both rotations and boosts. As a notional trick, one can define four-vectors A_α (subscript rather than superscript) as

$$A_\alpha \equiv g_{\alpha\beta} A^\beta, \tag{1.23}$$

which means that $A^\alpha A_\alpha$ is an invariant (summation of repeated indices inferred). In all calculations the summed indices always appear with one superscript (covariant) and one subscripted (contravariant). The fact that multiplying by $g_{\alpha\beta}$ simply lowers the index also applies to tensors, i.e.,

$$g_{\alpha\beta} C^{\beta\gamma} = C_\alpha^{\ \gamma}. \tag{1.24}$$

Note that this property means that g_β^α is simply the unit matrix. Furthermore, this property also applies to $g^{\alpha\beta}$,

$$g^{\alpha\beta} C_{\beta\gamma} = C^\alpha_{\ \gamma}. \tag{1.25}$$

Derivatives might at first seem a little backwards. Consider a scalar function $\phi(x)$, where x is a four vector. For small changes in x,

$$\delta\phi = \frac{\partial \phi}{\partial x^\mu} \delta x^\mu. \tag{1.26}$$

Since $\delta\phi$ is also a scalar, the vector $\partial/\partial x^\mu$ must transform as a contravariant vector. This motivates the notation,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \tag{1.27}$$

Note that for this convention, the equation of continuity is especially compact. For a four vector j, where j^0 refers to the charge density and j^i signals the current density, $\partial \cdot j = 0 \ \rightarrow \ \partial_t j_0 +$

$\boldsymbol{\nabla \cdot \vec{j} = 0}$. Maxwell equations also take on particularly beautiful forms, $\boldsymbol{\partial_\alpha F^{\alpha\beta} = j^\beta}$, and $\boldsymbol{\partial_\alpha \tilde{F}^{\alpha\beta} = 0}$. Additionally, we point out that we will employ the convention throughout this course that greek indices refer to all four components, while roman indices suggest only spatial components. Bold face will refer to the vector components of a three-vector, while four-vectors will not be put into bold face, i.e., $\boldsymbol{p \cdot x = p_0 x_0 - p \cdot x}$.

1.4 Four-Velocities and Momenta

For a particle that moves between two points by a small displacement $\boldsymbol{\Delta x^\alpha}$, one can define the quantity,

$$\boldsymbol{\Delta\tau \equiv \sqrt{\Delta x_\alpha \Delta x^\alpha}}, \tag{1.28}$$

which is an invariant. In the frame of the particle it is easy to see what this quantity represents, since in that frame the spatial components $\boldsymbol{\Delta x_i}$ are all zero. Thus $\boldsymbol{\Delta\tau}$ is the amount a clock, moving with the particle, has progressed during the differential displacement. Further, one can define a vector,

$$\boldsymbol{u^\alpha \equiv \frac{\Delta x^\alpha}{\Delta\tau}}, \tag{1.29}$$

which is also a four-vector since $\boldsymbol{\Delta x^\alpha}$ is a four-vector and $\boldsymbol{\Delta\tau}$ is a Lorentz scalar (or invariant). The four-vector $\boldsymbol{u^\alpha}$ is often referred to as the relativistic velocity, and given that $\boldsymbol{\Delta x^0 = \gamma \Delta\tau}$, one can see that

$$\boldsymbol{u^0 = \gamma,\ u^i = \gamma \frac{dx^i}{dt} = \gamma v^i.} \tag{1.30}$$

It is easy to see that

$$\boldsymbol{u^\alpha u_\alpha = 1.} \tag{1.31}$$

The momentum is defined by multiplying $\boldsymbol{u^\alpha}$ by the particle's mass $\boldsymbol{m}$, which is also a scalar. This then gives,

$$\boldsymbol{p^\alpha p_\alpha = m^2.} \tag{1.32}$$

The zeroth component of the momentum is identified as the energy. This gives the relation that for a particle at rest,

$$\boldsymbol{E = p^0 = m.} \tag{1.33}$$

Of course, the more famous relation, $\boldsymbol{E = mc^2}$, requires keeping track of all the factors of $\boldsymbol{c}$.

EXAMPLE:
A beam of particles of mass $\boldsymbol{m_A}$ is aimed a target with particles of mass $\boldsymbol{m_B}$. What kinetic energy, $\boldsymbol{K}$, is required so that one can make a resonance of mass $\boldsymbol{m_C}$.

The solution is based on energy-momentum conservation. One way to move forward is to calculate the invariant mass in terms of the total momentum and set it to $\boldsymbol{m_C}$.

$$\begin{aligned}
\boldsymbol{m^2_{\text{inv}}} &= \boldsymbol{(p_A + p_B)^2 = (K + m_A + m_B)^2 - \vec{p}_A^{\,2}} \\
&= \boldsymbol{(K + m_A + m_B)^2 - \left[(K + m_A)^2 - m_A^2\right]} \\
&= \boldsymbol{m_A^2 + m_B^2 + 2m_B(K + m_A) = m_C^2,} \\
\boldsymbol{K} &= \boldsymbol{\frac{m_C^2 - (m_A + m_B)^2}{2m_B}.}
\end{aligned}$$

EXAMPLE:
Consider two particles recorded with momenta p_A and p_B at space times separated by $r^\alpha = x_1^\alpha - x_2^\alpha$. In terms of relativistic invariants using p_A, p_B and r, solve for the impact parameter, i.e., the distance of closest approach as measured by an observer in the center-of-mass frame.

First, we consider two four vectors which will be used as projectors to eliminate the components of r that, in the center-of-mass frame, are either time-like or along the direction of the relative momentum. First, the time-like vector is the total momentum,

$$P^\alpha = p_A^\alpha + p_B^\alpha,$$

which in the c.o.m. frame becomes $(E_A + E_B, 0, 0, 0)$. One can define a vector,

$$r'^\alpha \equiv r^\alpha - P^\alpha \frac{P \cdot r}{P^2}.$$

In the center-of-mass frame, r' looks exactly like r, except the $\alpha = 0$ component vanishes. Next, one defines a vector which in the c.o.m. frame is parallel to the relative momentum, and is zero for the $\alpha = 0$ piece. This would be:

$$q'^\alpha = q^\alpha - P^\alpha \frac{P \cdot q}{P^2}, \;\; q^\alpha = p_A^\alpha - p_B^\alpha.$$

Again, one can consider the c.o.m. frame, where it is clear that $q'^0 = 0$ and $q'^i = (p_A^i - p_B^i)$. One can then project away the part of r' parallel to q, and make a new vector b,

$$b^\alpha \equiv r'^\alpha - q'^\alpha \frac{q' \cdot r'}{q'^2}$$

Finally, the impact parameter squared is, after some surprisingly painful algebra,

$$B^2 = -b^2 = -r^2 + \frac{(q \cdot r)^2 P^2 + (P \cdot r)^2 q^2 - 2(q \cdot r)(P \cdot r)(P \cdot q)}{P^2 q^2 - (P \cdot q)^2}.$$

1.5 Examples of Invariants

Here I discuss several invariants you will encounter at various times throughout this course or in the literature. First, one can see that d^4x is invariant to boosts by considering the Jacobian of a Lorentz transformation along the x axis. In that case,

$$\mathcal{L} = \begin{pmatrix} \cosh\eta & \sinh\eta & 0 & 0 \\ \sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.34}$$

The Jacobian is simply the determinant of the matrix, $J = \cosh^2 - \sinh^2 = 1$. Thus, d^4x is invariant. For a boost in an arbitrary direction, one can first to a rotation (d^3x is invariant to rotation) combined with a boost.
Then d^4p is also invariant as is,

$$\int d^4p\, \delta(p^2 - m^2) s(p) = \int dp_0 d^3p\, \delta(p_0^2 - E_p^2) s(p) = \int \frac{d^3p}{2E_p}\ s(p)|_{p_0=E_p}\,. \tag{1.35}$$

Here $s(p)$ is some arbitrary scalar function. Since s could be anything, then one can state that d^3p/E_p is also invariant. This is why spectra in particle physics are expressed as one of the two identical forms below

$$\frac{E_p dN}{d^3p} = \frac{dN}{d\phi p_t dp_t dy}. \tag{1.36}$$

The expression on the right-hand side can be found by seeing that $dp_z/E = dy$, where y is the rapidity. Since dN is the number of counts in a bin, and since a number of counts is invariant, you needn't worry about how dN transforms.
Often, when calculating the number of particles in a box, one writes

$$N = \int \frac{d^3x d^3p}{(2\pi)^3}\, f(p), \tag{1.37}$$

where f is the phase space density, or occupation probability. This must be an invariant. For instance, if you had fermions as zero temperature, f would be zero or unity depending on whether it was inside or outside the Fermi sea. This would not change if viewed from a different reference frame. Since f is invariant, and dN is invariant, it stands that $d^3x\ d^3p$ is also invariant. To see that this is true, one must remember that for this expression one assumes that dx measures the distance between the boundaries of a cell where both the positions of the boundaries are measured simultaneously, like measuring the length of a meter stick. If one compares the volume measured by an observer moving with the velocity of the particle, p/E, and compares to the length in the lab frame one finds that d^3x shorter by the Lorentz factor, E/m. If one considers d^3p in the two frames, one sees that d^3p in the lab frame must be larger by a factor of E/m so that d^3p/E is invariant. Thus, the product d^3pd^3x is invariant.
Another example that comes up often is a collision rate per volume and per time, dN_c/d^4x, which is manifestly invariant. One can also consider the collision rate of particles from specific regions of momentum space, d^3p_a/E_a and d^3p/E_b. The quantity would normally written as

$$E_a E_b \frac{dN_c}{d^4x\ d^3p_a\ d^3p_b} = f_a(p_a) f_b(p_b) Something, \tag{1.38}$$

where f_a and f_b are the phase space densities, or occupancies, of the two particles. Occupancies are also invariant, i.e. if the Fermi sea is full, $f = 1$ regardless of what frame is considered. Here, $Something$ has to be an invariant. Since it is invariant one can consider a frame where the momenta are back to back, $p_a = -p_b$. In that frame one knows from kinematics that

$$Something = \frac{E_a E_b}{(2\pi)^6} \sigma(\sqrt{s}) v_{\rm rel}, \tag{1.39}$$

where σ is the cross section and the relative velocity is $v_{\rm rel} = |p|/E_a + |p|/E_b$. One must then simply write $E_a E_b v_{\rm rel}$ as a Lorentz invariant. To do this,

$$E_a E_b v_{\rm rel} = |p_a E_b - p_b E_a|. \tag{1.40}$$

One can now write $E_a = p_a \cdot P/\sqrt{s}$ and replace p_a with the spatial components of $p'_a \equiv p_a - P(p_a \cdot P)/s$ to project out the zero$^{\rm th}$ components of p_a. Doing the same for p_b one then finds

$$\begin{aligned} (E_a E_b v_{\rm rel})^2 &= -\left[(p_b \cdot P)(p_a - P(p_a \cdot P)/s) - (p_a \cdot P)(p_b - P(p_b \cdot P)/s)\right]^2/s \\ &= -\left[(p_b \cdot P)p_a - (p_a \cdot P)p_b\right]^2/s \\ &= (p_a \cdot p_b)^2 - m_a^2 m_b^2. \end{aligned} \tag{1.41}$$

Here, numerous steps are omitted from the last step as this is related to a homework problem at the beginning of the last chapter. Putting this all together, one finds

$$\frac{dN_c}{d^4x} = \int \frac{d^3p_a}{(2\pi)^3 E_a}\frac{d^3p_b}{(2\pi)^3 E_b} f_a(p_a) f_b(p_b) \sigma(\sqrt{s}) \sqrt{(p_a \cdot p_b)^2 - m_a^2 m_b^2}. \tag{1.42}$$

1.6 Homework Problems

1. Suppose you are doing a fixed-target experiment at the LHC. The protons have a beam energy of 7 TeV. (The proton mass is 938.28 MeV/c^2). If the experiment were redone with a collider built with an equivalent center-of-mass energy, what would that energy be.

2. Find the equivalent fixed beam energy for a fixed target to have the same center-of-mass energy as the collider experiment at the LHC.

3. Consider a 1+1 dimension vector, (E, p), where $m^2 \equiv E^2 - p^2$. Consider the transformed vector, $p'_\alpha = L_{\alpha\beta} p^\beta$, where L is defined according to Eq. (1.19). Show that $m'^2 \equiv E'^2 - p'^2 = m^2$.

4. Consider two particles with four momenta p_a and p_b. Particle a is recorded at the space time point $r_a = (0, 0, 0, 0)$ and particle b is recorded at $r_b = r$. For an observer moving with particle a find the time at which particle b passes at the point of closest approach. Express your answer in terms of Lorentz invariants, i.e., dot products involving p_a, p_b and r.

5. Consider two particles of mass m_a and m_b with four momenta p_a and p_b. In terms of m_a, m_b and $p_a \cdot p_b$, find the relative velocity $|\vec{v}_a - \vec{v}_b|$ according to an observer in the center-of-mass frame.

6. The Lorentz transformation is a tensor, $\mathcal{L}^{\alpha\beta}$, which transforms some four vector p^α observed by an observer moving with four velocity u^α to a vector p'^α as determined by an observer moving with four-velocity u'^α.

$$\mathcal{L}^\alpha_{\ \beta} p^\beta = p'^\alpha.$$

Since $\mathcal{L}$ is a tensor it must be of the form,

$$\mathcal{L}^{\alpha\beta} = Au^{\alpha}u'^{\beta} + Bu'^{\alpha}u^{\beta} + Cu^{\alpha}u^{\beta} + Du'^{\alpha}u'^{\beta} + Eg^{\alpha\beta}, \tag{1.43}$$

where $A - E$ are scalar functions of u and u'. Since $u^2 = u'^2 = 1$, the only scalar function available is $u \cdot u'$. Consider the transformation from the rest frame $u = (1, 0, 0, 0)$ to the frame $u' = (\gamma, \gamma v, 0, 0)$. You know that the Lorentz transformation is:

$$\mathcal{L}^{\alpha}_{\ \beta} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Further, $u \cdot u' = \gamma$. From the form in Eq. (1.43),

$$\mathcal{L}^{\alpha}_{\ \beta} = \begin{pmatrix} A\gamma + B\gamma + C + D\gamma^2 + E & -A\gamma v - D\gamma^2 v & 0 & 0 \\ B\gamma v + D\gamma^2 v & -D\gamma^2 v^2 + E & 0 & 0 \\ 0 & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix}.$$

(a) Solve for the coefficients A through E in terms of $\gamma = u \cdot u'$.

(b) Show that for the four vector u',

$$\mathcal{L}^{\alpha\beta}u'_{\beta} = u^{\alpha}.$$

(c) Show that

$$\mathcal{L}^{\alpha\beta}(u, u')_{\beta\gamma}(u', u) = g^{\alpha}_{\ \gamma}.$$

2 Klein-Gordon Equation

2.1 Motivating the form of the Equation

The Schroedinger equation is often motivated by beginning with the non-relativistic expression for energy,

$$\boldsymbol{E = \frac{p^2}{2m} + V(r)}, \tag{2.1}$$

then using the substitutions, $\boldsymbol{p_i = -i\hbar\partial_i}$, and $\boldsymbol{H = i\hbar\partial_t}$, we get

$$\boldsymbol{i\hbar\partial_t\psi(r,t) = -\frac{\hbar^2}{2m}\nabla^2\psi(r,t) + V(r)\psi(r,t).} \tag{2.2}$$

Similarly, one can start with the relativistic equation for energy momentum,

$$\boldsymbol{E^2 = p^2 + m^2}, \tag{2.3}$$

and write down (setting $\hbar = 1$ and $\boldsymbol{c = 1}$)

$$\boldsymbol{-\partial_t^2\phi(r,t) + \nabla^2\phi(r,t) - m^2\phi(r,t) = 0.} \tag{2.4}$$

The potential is missing at this point. Relativistically, potentials can enter either by having a position-dependent mass, by adding a vector potential $\boldsymbol{A^\alpha}$ through minimal substitution, or both, and find the Klein-Gordon equation,

$$\boldsymbol{(i\partial_t - eA_0(r,t))^2\phi - (-i\partial_i - eA_i(r,t))^2\phi - (m + U(r,t))^2\phi = 0.} \tag{2.5}$$

or more covariantly,

$$\boldsymbol{(i\partial - eA)^2\phi - (m + U(r,t))^2\phi = 0.} \tag{2.6}$$

Note that here we have invoked the convention that squaring a four vector $\boldsymbol{Y}$ implies, $\boldsymbol{Y^2 = y_\alpha y^\alpha}$, which means that $\boldsymbol{\partial^2 = \partial_t^2 - \nabla^2}$. Electro-magnetic interactions are given by vector potentials, whereas $\boldsymbol{U}$ is a scalar potential. Later on, we will see that a scalar potential behaves differently than the zeroth component of a vector potential once one gets away from the non-relativistic limit.

For a complex field ϕ, one can identify a current density (for the moment with no potentials),

$$\boldsymbol{j^\alpha = \frac{i}{2}\left[(\partial^\alpha\phi^*)\phi - \phi^*\partial^\alpha\phi\right].} \tag{2.7}$$

One can see that $\boldsymbol{j}$ obeys a continuity equation,

$$\begin{aligned} \boldsymbol{\partial_\alpha j^\alpha} &= \boldsymbol{\frac{i}{2}\left[(\partial^2\phi^*)\phi - \phi^*\partial^2\phi\right]} \\ &= \boldsymbol{\frac{i}{2}\left[(m^2\phi^*)\phi - \phi^* m^2\phi\right] = 0.} \end{aligned} \tag{2.8}$$

2.2 Plane-Wave Solutions and Negative-Energy Solutions

First, we consider the solutions with no potential. Looking at eigenstates of the translation operator, ∂_i, i.e., $\phi(r,t) \sim e^{ip\cdot r}$. Solutions to the Klein-Gordon equation are then

$$\begin{aligned} \phi_p(r,t) &= e^{-ip\cdot x},\ p=(p_0,p_1,p_2,p_3),\ x=(t,x_1,x_2,x_3), \\ p_0 = \pm E_p &= \pm\sqrt{p^2+m^2}. \end{aligned} \tag{2.9}$$

The negative-energy solutions are difficult to interpret without going into the language of field theory, i.e., the negative energy solutions will correspond to removal of a particle.
For each of these solutions, one can identify a current density using Eq. (2.7). Consider the solutions,

$$\phi = e^{-ip\cdot x}/\sqrt{VE_p}. \tag{2.10}$$

The factor of $E_p^{-1/2}$ is chosen to give j the dimensions of density.

$$j^0 = \frac{p_0}{E_pV},\quad j = \frac{p}{E_pV}. \tag{2.11}$$

For $p_0 = E_p$, this solution is the familiar one for the density of particles of charge $+1$ moving with velocity p/E_p. However, for $p_0 = -E_p$ it appears odd as j_0 is negative. Again, this will only make sense once put into the context of field theory.

EXAMPLE:
Calculate the reflection and transmission coefficients for an incident plane wave of a complex scalar field with momentum p_x in the positive x direction,

$$m(x) = \begin{cases} M_A, & x<0 \\ M_B, & x>0 \end{cases}$$

Also show that the current is conserved.

> Since the K.G. equation involves ∂_x^2, the derivative $\partial_x\phi$ must be continuous as well as the wave function. This condition can also be generated from momentum conservation. Assuming a form,
>
> $$\psi(x) = \begin{cases} e^{ip_xx} + Be^{-ip_xx}, & x<0 \\ Ce^{ip'_xx}, & x>0 \end{cases}$$
>
> where $p'^2_x = p_x^2 + M_A^2 - M_B^2$. The boundary conditions demand
>
> $$A + B = C, \text{ and } -ip_xA + ip_xB = -ip'_xC,$$
>
> with $A=1$. The solutions for the reflection and transmission coefficients are:
>
> $$|B|^2 = \frac{(p_x-p'_x)^2}{(p_x+p'_x)^2},\ |C|^2 = \frac{4p_x^2}{(p_x+p'_x)^2}$$

Finally, the current densities are proportional to the momenta, i.e., $j_x = -i\phi^* \partial_x \phi$. Checking the currents on the left and right,

$$\begin{aligned} p_x \left(1 - \frac{(p_x - p'_x)^2}{(p_x + p'_x)^2}\right) &=? \quad p'_x \frac{4p_x^2}{(p_x + p'_x)^2} \\ p_x \left(\frac{4p_x p'_x}{(p_x + p'_x)^2}\right) &=? \quad p'_x \frac{4p_x^2}{(p_x + p'_x)^2} \quad \checkmark \end{aligned}$$

2.3 Beginning from the Lagrangian

It is usually the Lagrangian (or more accurately the Lagrangian density) that is considered the foundation of field theories. The Klein-Gordon equation and the Dirac equation are then the equations of motion derived from the Lagrange equations of motion. In classical mechanics, one has one or more generalized coordinates q, and given $L(q, \dot{q})$, the equations of motion are

$$\frac{d}{dt}\frac{\partial L(q,\dot{q})}{\partial \dot{q}} - \frac{\partial L(q,\dot{q})}{\partial q} = 0. \tag{2.12}$$

For field theories, one considers the Lagrangian density rather than the Lagrangian, which means that the action is

$$S = \int d^4x \, \mathcal{L}(\phi(x), \partial_\mu \phi(x)). \tag{2.13}$$

The variation of the action is then

$$\begin{aligned} \delta S &= 0 \\ &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial_\mu \delta\phi(x) + \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi(x) \right\} \\ &= -\int d^4x \left\{ \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} \right\} \delta\phi(x), \end{aligned} \tag{2.14}$$

where the last step involved an integration by parts. Since the variation should go to zero for all variations $\delta\phi$ at any point x,

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{2.15}$$

These are Lagrange's equations of motion for fields. If the field has multiple components, one need only add an index j, so that $\phi \to \phi_j$, which leads to an equation of motion for each component. For a complex field, one would have two components, the real and imaginary part. Rather than deriving and writing equations of motion for the real and imaginary part separately, it is easier to derive the equations of motion by rewriting,

$$\begin{aligned} \delta S &= -\int d^4x \left\{ \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} \right] \delta\phi(x) + \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^*} - \frac{\partial \mathcal{L}}{\partial \phi^*} \right] \delta\phi^*(x) \right\} \\ &= -\int d^4x \left\{ \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} \right] + \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^*} - \frac{\partial \mathcal{L}}{\partial \phi^*} \right] \right\} \delta\phi_r \\ &\quad -\int d^4x \left\{ \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} \right] - \left[\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^*} - \frac{\partial \mathcal{L}}{\partial \phi^*} \right] \right\} \delta\phi_i \end{aligned} \tag{2.16}$$

If $\mathcal{L}$ is a real function of ϕ, one can express the equations of motion as a single equation (effectively two equations since it has a real and imaginary part):

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^*} - \frac{\partial \mathcal{L}}{\partial \phi^*} = 0. \tag{2.17}$$

The Lagrangian density for the K.G. equation of a complex scalar field interacting with an electromagnetic field is:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\phi^*(x)(i\partial_\mu - eA_\mu)(i\partial^\mu - eA^\mu)\phi(x) - \frac{1}{2}m^2\phi^*(x)\phi(x) \\ &= \frac{1}{2}\left[(-i\partial^\mu - eA^\mu)\phi^*\right]\left[(i\partial_\mu - eA_\mu)\phi\right] - \frac{1}{2}m^2\phi^*(x)\phi(x), \end{aligned} \tag{2.18}$$

where the last step involved integrating by parts to make the derivative act on the l.h.s. so that the Lagrangian is a manifestly real function of ϕ.
The Lagrangian equations of motion are:

$$\begin{aligned} \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^*} &= \frac{\partial \mathcal{L}}{\partial \phi^*} \\ \partial_\mu \left\{ \frac{-i}{2}(i\partial^\mu - eA^\mu)\phi \right\} &= -\frac{m^2}{2}\phi - \frac{eA^\mu}{2}(i\partial_\mu - eA_\mu)\phi. \end{aligned} \tag{2.19}$$

Moving everything to the l.h.s., one gets the K.G. equation.
One can calculate the current by considering the $\phi = \phi_{\mathbf{mag}} e^{i\psi}$, and noticing that the Lagrangian depends only on $\partial_\mu\psi$, not on ψ. Noether's theorem is enacted by solving for the equations of motion for ψ,

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = 0, \tag{2.20}$$

to find the current,

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = \mathbf{Re}\left[\phi^*(i\partial^\mu - eA^\mu)\phi\right]. \tag{2.21}$$

2.4 Lagrangian for Vector Field

For a vector field, A^μ, one can write a Lagrangian,

$$\mathcal{L} = \frac{-1}{4}F^{\mu\nu}F_{\mu\nu} + ej^\mu A_\mu, \;\; F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{2.22}$$

The field tensor is anti-symmetric and has six components,

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}, \tag{2.23}$$

where we have defined $\vec{B} = \nabla \times \vec{A}$, $\vec{E} = \partial_t \vec{A} - \nabla A_0$.

The equations of motion from the Lagrangian are

$$\begin{aligned} \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu A_\nu} &= \frac{\partial \mathcal{L}}{\partial A_\nu} \\ \frac{1}{2}\partial_\mu \left(F_{\mu\nu} - F_{\nu\mu}\right) &= j^\nu, \\ \partial_\mu F^{\mu\nu} &= j^\nu. \end{aligned} \tag{2.24}$$

This last equation are some of Maxwell's equations. As a HW problem, you will rewrite these in terms of $\vec{E}$ and $\vec{B}$. The remainder of Maxwell's equations come from the fact that

$$\begin{aligned} \partial_\mu \tilde{F}^{\mu\nu} &= 0, \\ \tilde{F}^{\mu\nu} &\equiv \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}. \end{aligned} \tag{2.25}$$

The right-hand-side is zero because

$$\begin{aligned} \partial_\mu \tilde{F^{\mu\nu}} &= \partial_\mu \epsilon^{\mu\nu\alpha\beta} \partial_\alpha A_\beta, & (2.26) \\ &= \epsilon^{\mu\nu\alpha\beta} \partial_\mu \partial_\alpha A_\beta \\ &= 0, & (2.27) \end{aligned}$$

where the last step used the fact that the anti-symmetry of $\epsilon^{\mu\nu\alpha\beta}$ will give zero when contracted with $\partial_\mu \partial_\alpha$. Your HW problem will also have you write this in terms of $\vec{E}$ and $\vec{B}$.

2.5 Homework Problems

1. Consider the case for the complex scalar field interacting with a vector field, see Eq. (2.5). Find an expression for the current, j^α, in terms of ϕ, derivatives and eA, such that one satisfies the equation of continuity,
$$\partial_\alpha j^\alpha = 0.$$

2. Consider the case of a particle of momentum $p_x > 0$ incident on a step function potential as was done in the example above, only with the potential being the zeroth component of a vector potential,
$$eA_0(x) = \begin{cases} 0, & x < 0 \\ V_0, & x > 0 \end{cases}$$
Let the mass, m, be constant and let $A_i(x) = 0$ everywhere.

   (a) Calculate the transmission and reflection coefficients.

   (b) Show that current is conserved across the boundary.

   (c) What range(s) of V_0 allow for propagating solutions (not exponentially damped) for $x > 0$?

   (d) Calculate and plot both the charge density, j_0, and current density, j_x, as a function of x for one example in each separate range in (c). Be sure to plot for both positive and negative x.

3. Consider the Lagrangian density for the electro-magnetic field, $\boldsymbol{A^\mu(x)}$, coupled to an external current $\boldsymbol{j(x)}$.

$$\boldsymbol{\mathcal{L} = \frac{-1}{4} F^{\mu\nu} F_{\mu\nu} + ej \cdot A, \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu.}$$

   (a) Solve for the equations of motion for $\boldsymbol{A^\mu}$, and show that this is equivalent to four of Maxwell's equations. (Derive and express the four equations in terms of $\vec{\boldsymbol{E}}$ and $\vec{\boldsymbol{B}}$).

   (b) The other 4 equations come from the identity $\boldsymbol{\partial_\alpha \tilde{F}^{\alpha\beta} = 0}$, where $\boldsymbol{\tilde{F}^{\alpha\beta} \equiv \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta}}$. Since $\boldsymbol{\epsilon^{\alpha\beta\gamma\delta}}$ is anti-symmetric in the indices, and since $\boldsymbol{\partial_\mu}$ appears twice in $\boldsymbol{\partial_\alpha \tilde{F}^{\alpha\beta}}$ (once in the definition of $\boldsymbol{F^{\mu\nu}}$),

$$\begin{aligned} \boldsymbol{\partial_\alpha \tilde{F}^{\alpha\beta}} &= \boldsymbol{\epsilon^{\alpha\beta\gamma\delta} \partial_\alpha (\partial_\gamma A_\delta - \partial_\delta A_\gamma)} \\ &= \boldsymbol{0.} \end{aligned} \tag{2.28}$$

   Express these equations, $\boldsymbol{\partial_\alpha \tilde{F}^{\alpha\beta} = 0}$, in terms of $\vec{\boldsymbol{E}}$ and $\vec{\boldsymbol{B}}$.

3 Dirac Equation

3.1 Motivating the form of the equation, and γ matrices

Dirac's idea was to create an equation that is linear in the energy and momentum, and is consistent with $E^2 = p^2 + m^2$. For spin 1/2 particles one is accustomed to the wave functions being a spinor, i.e.,

$$\Psi(x) \rightarrow \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}, \tag{3.1}$$

where the two components are related to the amplitudes for observing a spin-up or spin-down particle. Further, some elements of the Hamiltonian might mix the upper and lower components, such as magnetic coupling to a magnetic field in the x or y directions.
For the Dirac equation, we will also consider the wave function to be a spinor, but for the moment will not specify the dimension of the spinor. The Hamiltonian can be a matrix. We will assume,

$$H = P_x\alpha_x + P_y\alpha_y + P_z\alpha_z + m\beta, \tag{3.2}$$

where P_i are the momentum operators, m is the mass, and α_i and β are matrices. Squaring the Hamiltonian,

$$\begin{aligned} H^2 &= P_x^2\alpha_x^2 + P_y^2\alpha_y^2 + P_z^2\alpha_z^2 + m^2\beta^2 \\ &+ P_xP_y\{\alpha_x,\alpha_y\} + P_xP_z\{\alpha_x,\alpha_z\} + P_yP_z\{\alpha_y,\alpha_z\} + mP_x\{\beta,\alpha_x\} + mP_y\{\beta,\alpha_y\} + mP_z\{\beta,\alpha \end{aligned}$$

In order for $E^2 = |\vec{P}|^2 + m^2$ to be true, the four matrices, $\alpha_x, \alpha_y, \alpha_z, \beta$, must square to unity and must mutually anti-commute. If one needed only three matrices, one could work in two dimensions and use the three σ matrices. The spinors could thus have two dimensions. Oddly enough, there are no sets of three-dimensional matrices that work. Instead, one must go to four-dimensional matrices. There are innumerable representations of such matrices. The most common is known as the Dirac representation,

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}. \tag{3.4}$$

Another notation, which we will motivate later, are the γ matrices,

$$\gamma^i \equiv \beta\alpha^i, \;\; \gamma^0 \equiv \beta. \tag{3.5}$$

The commutation relations for the γ matrices are:

$$\{\gamma^\alpha, \gamma^\beta\} = 2g^{\alpha\beta}\mathbb{I}. \tag{3.6}$$

For the Dirac representation,

$$\gamma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}. \tag{3.7}$$

Stating H as $i\partial_t$, the Dirac equation can be written in terms of the γ matrices (simply multiply by γ^0 as:

$$(i\gamma^\mu\partial_\mu - m)\psi = 0, \tag{3.8}$$

or using the Feynmann "slash" notation , $\not{a} \equiv a^\mu \gamma_\mu$,

$$i\not{\partial}\psi - m\psi = 0. \tag{3.9}$$

EXAMPLE:
Calculate $\mathbf{Tr}\not{\gamma}$,

From the anti-commutation relations,

$$\not{\gamma} = \gamma^\alpha \gamma_\alpha = (1/2)\{\gamma^\alpha, \gamma_\alpha\} = g^\alpha_{\ \alpha}\mathbb{I}.$$

Taking the trace and summing over α,

$$\sum_\alpha g^\alpha_{\ \alpha} \mathrm{Tr}\ \mathbb{I} = 4\sum_\alpha = 16.$$

3.2 Plane-wave solutions

Consider solutions of the form $e^{-iE_p t + ip\cdot x}u(p)$, where $u(p)$ is a 4-component matrix. If we use ϕ and χ to refer to the upper two and lower-two components separately,

$$u(p) = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{3.10}$$

then using the Dirac representation for the γ matrices,

$$\begin{aligned} E_p\phi &= m\phi + \sigma\cdot p\chi, \\ -E_p\chi &= m\chi - \sigma\cdot p\phi. \end{aligned} \tag{3.11}$$

One can now guess at a solution with ϕ being an eigenstate of $\sigma\cdot p$,

$$(\sigma\cdot p)\phi = \pm|p|\phi. \tag{3.12}$$

Inserting into the lower expression in Eq. (3.11),

$$\chi = \pm\frac{|p|}{E_p + m}\phi. \tag{3.13}$$

Inserting back into the upper equation to test for consistency,

$$(E_p - m)\phi = \sigma\cdot p\left[\pm\frac{|p|}{E_p+m}\phi\right] = \frac{|p|^2}{E_p+m}\phi, \tag{3.14}$$

which checks out since $(E_p^2 - m^2) = |p|^2$. Thus the lower two components are smaller than the upper two components by a factor $|p|/(E_p + m)$. For a particle at rest, $|p| = 0$, the lower components are zero.

One could have also looked for solutions with negative energy, i.e., plane waves of the form $e^{iE_pt+ip\cdot x}u(p)$, with E_p still referring here to the positive square root, $\sqrt{|p|^2+m^2}$. Repeating the method above, one would have found

$$\chi = \mp\frac{|p|}{E_p - m}\phi, \tag{3.15}$$

or equivalently,

$$\phi = \mp\frac{|p|}{E_p + m}. \tag{3.16}$$

Thus for the negative energy solutions, the lower components are larger. Usually, the solutions with negative energy are referred to as $v(p)$ rather than $u(p)$.
A second way to derive the solutions for ϕ and χ is to note that the Dirac equation, $(\not{p}-m)\psi = 0$ will be satisfied for

$$\psi = (\not{p} + m)\psi(p = 0), \tag{3.17}$$

because

$$(\not{p} - m)(\not{p} + m) = p^2 - m^2 = 0. \tag{3.18}$$

Since the Hamiltonian for zero momentum is $m\gamma_0$, the eigenstates ψ_0 have eigen-energies of $\pm m$, with the sign depending on whether ψ_0 is either confined completely to the upper or completely to the lower components. One can then rederive the expressions for $\phi(p)$ and $\chi(p)$ rather easily.

EXAMPLE:
Solve the one-dimensional square-well for the case where the mass is infinite outside the well,

$$m = \begin{cases} \infty, & x < 0 \\ 0, & 0 < x < L \\ \infty, & x > L \end{cases}$$

Unlike the Schrödinger equation, there are only first derivatives in the Dirac equation. Thus, though the wave function must be continuous, the derivative need not be. Further, for an infinite well the wave function need not go to zero at the boundary since the wave function can decay exponentially in effectively zero distance. Let the mass for $x > L$ be $M \to \infty$. The solution to the Dirac equation in the three regions is of the form

$$\psi(x) = \begin{cases} e^{\kappa x}u_1, & x < 0, & \kappa = \sqrt{M^2 - E^2} \\ Ae^{ikx}u_2 + Be^{-ikx}u_3, & 0 < x < L, & k = E \\ e^{-\kappa(x-L)}u_4, & x > L, & \kappa = \sqrt{M^2 - E^2}. \end{cases}$$

To solve for the BC, we consider the Dirac equation

$$E\psi = -i\alpha_x\partial_x\psi + M\beta\psi.$$

Since $M \to \infty$, we can set the left-hand side to zero. Integrating the Dirac equation from $x = L$ to ∞ using the fact that $\psi \sim e^{-M(x-L)}$ for $x > L$, gives

$$0 = i\alpha_x\psi(L) + \beta\psi(L).$$

This gives the boundary condition,

$$-i\gamma_x\psi(x=L)=\psi(x=L).$$

For $x=0$, the one integrates the Dirac equation from $-\infty$ to 0 to obtain the second BC,

$$i\gamma_x\psi(x=0)=\psi(x=0).$$

For arbitrary directions, one can write BC for massless particles at a boundary with an infinite scalar field as,

$$i\vec{\gamma}\cdot\hat{n}\psi=\psi, \tag{3.19}$$

where $\hat{n}$ points into the well. Thus, there are two BC but not the ones one would have expected. To proceed we note that the solutions for left and right-going waves are related by the γ_x matrix,

$$\gamma_x u_2=u_3,\ \gamma_x u_3=-u_2.$$

To verify this consider the operator $p_x\alpha_x$ operating on $\gamma_x u_2$. Since γ_x anti-commutes with α_x, one can see that $p_x\alpha_x u_2=-Eu_2$, which is the positive energy solution for a particle with negative momentum.

Inserting the definition for ψ into the BC at $x=0$,

$$\begin{aligned} i\gamma_x\left(Au_2+Bu_3\right) &= Au_2+Bu_3, \\ iAu_3-iBu_2 &= -Au_2-Bu_3. \end{aligned}$$

This gives

$$A=iB.$$

One can now repeat for the BC at $x=L$.

$$\begin{aligned} i\gamma_x\left(Au_2+Bu_3e^{-2ikL}\right) &= -\left(Au_2+Bu_3e^{-2ikL}\right), \\ iAu_3-iBu_2e^{-2ikL} &= Au_2+Bu_3e^{-2ikL}. \end{aligned}$$

This gives

$$A=-iBe^{-2ikL}.$$

For both BC to be satisfied,

$$e^{-2ikL}=-1,\ \text{or}\ k=\frac{(2n+1)\pi}{2L},\ n=0,1,2,\cdots$$

Note that these are not the same values of k one would have obtained for the Schrödinger equation. Although the separations between levels are the same, the ground state's k is half of what one would have obtained.

3.3 Boosting and Rotating

One operator that commutes with the Dirac equation is

$$\boldsymbol{\sigma_{ij}} = \frac{\boldsymbol{i}}{\boldsymbol{2}}[\gamma_i, \gamma_j]. \tag{3.20}$$

This operator is Hermitian, and can be shown to rotate the state ψ when applied

$$\boldsymbol{R(\omega)\psi} = \boldsymbol{e^{-i\sigma_{ij}\omega_{ij}/4}\psi}. \tag{3.21}$$

More familiarly, one can define the Spin operator as

$$\boldsymbol{S_i} = \frac{\boldsymbol{i}}{\boldsymbol{2}}\boldsymbol{\epsilon_{ijk}\sigma_{jk}}, \tag{3.22}$$

which has eigenvalues of $\boldsymbol{\pm 1/2}$. The anti-symmetric matrix $\boldsymbol{\omega}$ can be thought of as representing the rotation angle, $\boldsymbol{\theta_i = \epsilon_{ijk}\omega_{jk}/2}$.
When observing the form of the Lorentz matrices, boosts appeared similar to rotations, with the rapidity $\boldsymbol{\eta}$ playing the role of the angle. Similarly, we will guess here that the boost matrix would have the same form, only with $\boldsymbol{\omega_{ij}}$ having one spatial and one time index. To that end, we consider the matrix,

$$\boldsymbol{S(\omega_{\alpha\beta})} \equiv \boldsymbol{e^{-i\sigma^{\alpha\beta}\omega_{\alpha\beta}/4}}. \tag{3.23}$$

The inverse matrix is

$$\boldsymbol{S^{-1}(\omega_{\alpha\beta})} = \boldsymbol{e^{i\sigma^{\alpha\beta}\omega_{\alpha\beta}/4}} \neq \boldsymbol{S^\dagger}. \tag{3.24}$$

The fact that $\boldsymbol{S^\dagger \neq S^{-1}}$ can be understood by noticing that $\boldsymbol{\sigma_{0i}}$ is not Hermittian. To see that $\boldsymbol{S}$ is the boost matrix, we consider the transformation of the γ matrices when $\boldsymbol{\omega_{\alpha\beta}}$ has only an element, $\boldsymbol{\omega_{0z} = -\omega_{z0} = \eta}$. The transformation then becomes

$$\begin{aligned}
\gamma^\alpha &\rightarrow \boldsymbol{S^{-1}\gamma^\alpha S} \\
\gamma^z &\rightarrow \boldsymbol{e^{\gamma^0\gamma^z\eta/2}\gamma^z e^{-\gamma^0\gamma^z\eta/2}} \\
&= \boldsymbol{\left[\cosh\eta/2 + \sinh(\eta/2)\gamma^0\gamma^z\right]\gamma^z\left[\cosh\eta/2 - \sinh(\eta/2)\gamma^0\gamma^z\right]} \\
&= \boldsymbol{\left[\cosh^2(\eta/2) + \sinh^2(\eta/2)\right]\gamma^z - 2\sinh(\eta/2)\cosh(\eta/2)\gamma^0} \\
&= \boldsymbol{(\cosh\eta)\gamma^z - (\sinh\eta)\gamma^0},
\end{aligned} \tag{3.25}$$

which given that $\boldsymbol{\cosh\eta = \gamma}$ is the Lorentz contraction factor, and $\boldsymbol{\sinh\eta = \gamma v}$, shows that $\boldsymbol{S}$ boosts the various γ^α just like it should for a Lorentz boost.
Since $\boldsymbol{S^{-1} \neq S^\dagger}$, the quantity $\psi^\dagger\psi$ is not a Lorentz invariant, i.e.,

$$\boldsymbol{\psi^\dagger S^\dagger S\psi} \neq \boldsymbol{\psi^\dagger\psi}. \tag{3.26}$$

However, it is easy to see that

$$\boldsymbol{\gamma^0 S^\dagger \gamma^0 = S^{-1}}. \tag{3.27}$$

To see this, remember that γ^0 is Hermitian while γ_i are anti-Hermitian. One can now see that the quantity $\psi^\dagger\gamma^0\psi$ is a Lorentz invariant, because

$$\begin{aligned}
\boldsymbol{\psi^\dagger S^\dagger\gamma^0 S\psi} &= \boldsymbol{\psi^\dagger\gamma^0\left(\gamma^0 S^\dagger\gamma^0\right)S\psi} \\
&= \boldsymbol{\psi^\dagger\gamma^0 S^{-1}S\psi = \psi^\dagger\gamma^0\psi}.
\end{aligned} \tag{3.28}$$

The quantity $\psi^\dagger\gamma^0$ occurs often and is referred to as

$$\bar{\psi} \equiv \psi^\dagger\gamma^0, \tag{3.29}$$

and for nearly all expressions one will see $\bar{\psi}$ used rather than expressions with $\psi^\dagger$. Thus, a Lorentz invariant is $\bar{\psi}\psi$.

EXAMPLE:
Show that for a plane wave solution, $\bar{\psi}\psi = (m/E)\psi^\dagger\psi$.

First consider that fact that $u(\mathrm{p}) \equiv (\not{p} + m)u(\mathrm{p} = 0)$ is a solution to the Dirac equation. The ratio in question is then

$$\begin{aligned}\frac{\bar{u}(\mathrm{p})u(\mathrm{p})}{u^\dagger(\mathrm{p})u(\mathrm{p})} &= \frac{u^\dagger(0)(\not{p}^\dagger + m)\gamma_0(\not{p} + m)u(0)}{u^\dagger(0)(\not{p}^\dagger + m)(\not{p} + m)u(0)} \\ &= \frac{u^\dagger(0)\gamma_0(\not{p} + m)(\not{p} + m)u(0)}{u^\dagger(0)(\not{p}^\dagger + m)(\not{p} + m)u(0)},\end{aligned}$$

where we have made use of the fact that $\gamma_i^\dagger\gamma_0 = \gamma_0\gamma_i$. For the final steps we use that fact the $\gamma_0 u(0) = u(0)$, and that only even powers of γ_i will contribute to any of the matrix elements, because the γ_i move one from upper to lower components while $u(0)$ is only in the upper components. These considerations yield

$$\begin{aligned}\frac{\bar{u}(\mathrm{p})u(\mathrm{p})}{u^\dagger(\mathrm{p})u(\mathrm{p})} &= \frac{(E^2 - \vec{p}^2 + m^2 + 2mE)u^\dagger(0)u(0)}{(E^2 + \vec{p}^2 + m^2 + 2mE)u^\dagger(0)u(0)} \\ &= \frac{2m^2 + 2mE}{2E^2 + 2mE} = \frac{m}{E}.\end{aligned}$$

3.4 Conserved Currents

Given that $\bar{\psi}\psi$ is a scalar, $\bar{\psi}\gamma^\mu\psi$ is a four vector with dimensions of density. This makes it a candidate for a current. To test for current conservation,

$$\begin{aligned}\partial_\mu(\bar{\psi}\gamma^\mu\psi) &= (\partial_\mu\psi^\dagger)\gamma^0\gamma^\mu\psi + \psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi \\ &= (\gamma^0\gamma^\mu\partial_\mu\psi)^\dagger)\psi + \psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi \\ &= -im\bar{\psi}\psi + im\bar{\psi}\psi = 0,\end{aligned} \tag{3.30}$$

where the second line was derived using the fact that $\gamma^0\gamma^\mu$ is Hermitian, and the Dirac equation was applied in the third line. It is noteworthy that the momentum operator was not involved in defining the current.
For a momentum state, one can the write the wavefunction as $\psi = e^{-ip\cdot x}u(p)$ with

$$u(p) = \frac{1}{Z^{1/2}(p)}(\not{p} + m)u(p = 0), \tag{3.31}$$

where $Z(p)$ is a normalization coefficient. Again, it easy to see that this is a solution of $(\not{p} - m)u(p) = 0$. The current is then,

$$\begin{aligned} j^\mu &= \frac{1}{Z}u^\dagger(p=0)(\not{p}^\dagger + m)\gamma^0\gamma^\mu(\not{p}+m)u(p=0) \\ &= \frac{1}{Z}u^\dagger(p=0)\gamma^0(\not{p}+m)\gamma^\mu(\not{p}+m)u(p=0). \end{aligned} \tag{3.32}$$

Commuting γ^μ through $\not{p} + m$ using the commutation relation, $\{\not{p}, \gamma^\mu\} = 2p^\mu$,

$$j^\mu = \frac{1}{Z}u^\dagger(p=0)\gamma^0\gamma^\mu(-\not{p}+m)(\not{p}+m)u(p=0) + \frac{2p^\mu}{Z}u^\dagger(p=0)\gamma^0(\not{p}+m)u(p=0). \tag{3.33}$$

The first term can be thrown away since $(\not{p} - m)(\not{p} + m) = 0$, and the second term is easy to evaluate in the Dirac representation since γ^μ mixes upper and lower components, while $u(p = 0)$ has only upper components. One then finds,

$$j^\mu = \frac{2p^\mu}{Z}u(p=0)^\dagger\gamma^0(p_0\gamma^0 + m)u(p=0) = \frac{2p^\mu}{Z}(E_p + m)\bar{u}(p=0)u(p=0). \tag{3.34}$$

The ratio of the spatial currents, j_i, to the charge density, j^0, is indeed

$$\frac{j_i}{j_0} = \frac{p_i}{E_p}, \tag{3.35}$$

which is the velocity. Thus j^μ indeed depends on p^μ in the way one would expect.

3.5 Beginning from the Lagrangian

The Dirac equation gives a particularly simple form for the Lagrangian density,

$$\mathcal{L} = \bar{\psi}(x)(i\not{\partial}\psi(x) - m\bar{\psi}(x)\psi(x). \tag{3.36}$$

With this form the action, $\int d^4x\mathcal{L}$, is clearly a Lorentz invariant. Starting from the Lagrangian, we will derive numerous items (many already presented thus far): the Dirac Equation, the stress-energy tensor, and the conserved charge density.
To generate the equations of motion for ψ, one uses the relation

$$\partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\psi_i)} = \frac{\partial\mathcal{L}}{\partial\psi_i}, \tag{3.37}$$

where the subscript i refers to either the real or imaginary part of ψ. Equivalently, one can write the two equations as

$$\partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\psi)} = \frac{\partial\mathcal{L}}{\partial\psi}, \tag{3.38}$$

with the implication that $\psi^\dagger$ is fixed. This gives the equation,

$$i\partial_\alpha\bar{\psi}\gamma^\alpha = -m\bar{\psi}, \tag{3.39}$$

which is the Dirac equation for $\bar{\psi}$. If you take the complex conjugate of the above equation, one obtains,

$$\begin{aligned} -i\gamma^0\gamma^\alpha\partial_\alpha\psi &= -m\gamma^0\psi, \\ i\not{\partial}\psi &= m\psi, \end{aligned} \tag{3.40}$$

which is the Dirac equation for ψ.
The conserved currents are generated by noting that the Lagrangian is invariant to changes of phase, $\psi \rightarrow e^{i\phi}\psi$, where ϕ is independent of x. Treating ϕ as a field, the equations for ϕ are

$$\partial_\alpha \frac{\partial \mathcal{L}}{\partial_\alpha \phi} = \frac{\partial \mathcal{L}}{\partial \phi}. \tag{3.41}$$

Since $\mathcal{L}$ doesn't depend on ϕ, only on $\partial_\alpha\phi$, the r.h.s. is zero and one obtains

$$\partial_\alpha \frac{\partial}{\partial \partial_\alpha \phi}\left(\bar{\psi}\gamma^\beta(i\partial_\beta\phi)\psi\right) = 0, \tag{3.42}$$

which shows that one can identify the conserved current as

$$j^\alpha = \bar{\psi}\gamma^\alpha\psi. \tag{3.43}$$

Finally, one can derive the stress-energy tensor,

$$\begin{aligned} T_{\alpha\beta} &= \frac{\partial \mathcal{L}}{\partial(\partial^\alpha\psi)}\partial_\beta\psi - g_{\alpha\beta}\mathcal{L} \\ T_{00} &= \bar{\psi}(i\vec{\gamma}\cdot\nabla + m)\psi, \\ T_{0i} &= i\bar{\psi}\gamma_0\partial_i\psi, \\ T_{i0} &= -i\bar{\psi}\gamma^i\partial_t\psi, \\ T_{ij,i\neq j} &= i\bar{\psi}\gamma_i\partial_j\psi \\ \sum_i T_{ii} &= \bar{\psi}(i\gamma^0\partial_t - m)\psi. \end{aligned} \tag{3.44}$$

If one wants to eliminate the time derivatives above, the Dirac equation can be invoked to replace $\partial_t\psi$. One problem with the tensor above is that it is not symmetric (See Landau, Classical Theory of Fields, sec. 32). To correct for that, one needs to add a term

$$\Delta T^{\alpha\beta} = -\frac{1}{2}\bar{\psi}(\gamma^\alpha P^\beta - P^\alpha\gamma^\beta)\psi. \tag{3.45}$$

Once can show that $\partial_\alpha \Delta T^{\alpha\beta} = 0$, and that by adding it to $T^{\alpha\beta}$, one obtains a symmetric expression for $T^{\alpha\beta}$. As a homework problem, it can be shown that $\partial_\alpha \Delta T^{\alpha\beta} = 0$.

3.6 Homework Problems

1. Consider the chiral representation of the Dirac matrices,

$$\beta = \begin{pmatrix} 0 & -\mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}, \quad \alpha^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}.$$

(a) Show that for this representation that $(p^i\alpha^i + m\beta)^2 = |p|^2 + m^2$.

(b) Find both solutions for $u(p)$ and for $v(p)$ for p along the positive z axis.

(c) Write the two solutions above for the limit $p \to \infty$ and for $p \to 0$.

(d) The spin operator, $\Sigma_k = (-i/2)\epsilon_{ijk}\gamma_i\gamma_j$ in this representation is

$$\Sigma_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}.$$

For massless particles, find $u(p)$ with positive energy and positive helicity, i.e., $(p \cdot \Sigma)u = +|p|u$. Then repeat for solutions with the other three combinations of positive/negative energies and helicities.

(e) For massless particles, one could get by with only two-by-two matrices. Comment on what helicities would be possible if the Dirac matrices were only two-by-two.

2. Show that $\not{p}^2 = p^2$.

3. Show that $v(-\vec{p}) = (\vec{\alpha} \cdot \hat{p})u(\vec{p})$. I.e., show that if u is a positive energy solution, that $(\vec{\alpha} \cdot \hat{p})u$ will be a negative energy solution.

4. Show that $(\not{p} + m)/2m$ is a projection operator that projects out either the positive or negative energy solutions for momentum p depending on whether p_0 is $+E_p$ or $-E_p$.

5. Consider the spinor $u(p) = \frac{1}{Z(p)}(\not{p} + m)u_0$, where $u_0 = (1, 0, 0, 0)$ is the $p = 0$ solution in the Dirac representation, and let p point in the z direction. Find $Z(p)$ such that $u = Su_0$, where the boost matrix S is defined

$$S = e^{\gamma^0\gamma^z\eta/2},$$

and η is the rapidity, i.e., $\sinh\eta = |\vec{p}|/m$.

6. Using Eq.s (3.43) and (3.44), calculate the ratio of the energy density to the charge density, T_{00}/j_0, and the momentum density to the charge current, T_{0i}/j_0, for a plane wave solution.

7. Consider the matrix
$$\gamma_5 \equiv i\gamma^0\gamma^x\gamma^y\gamma^z.$$
Independent of the representation,
   (a) Show that $\gamma_5^2 = 1$.
   (b) Show that $(1+\gamma_5)/2$ is a projector.
   (c) Show that $\gamma_5^\dagger = \gamma_5$.
   (d) Show that $\{\gamma^\alpha, \gamma_5\} = 0$.
   (e) Show that $\mathrm{Tr}\ \gamma_5 = 0$.
   (f) Given the properties above, show that γ_5 has two eigenvalues of -1 and two eigenvalues of $+1$.
   (g) Given that $\mathrm{Tr}\ \gamma^\alpha = 0$, show that $\mathrm{Tr}\ \gamma^\alpha\gamma^\beta\gamma^\delta = 0$ for any $\alpha\beta\gamma$.
   (h) Consider a positive-energy plane-wave solution for massless particles with positive/negative helicity,
$$\begin{aligned}(p\cdot\alpha)u_\pm(p) &= E_p u_\pm(p),\\ (p\cdot\Sigma)u_\pm(p) &= \pm E_p u_\pm(p),\end{aligned}$$
   where the spin operator $\Sigma_k = -i\epsilon_{ijk}\alpha_i\alpha_j/2$. Show that for such solutions,
$$\gamma_5 u_\pm(p) = \pm u_\pm(p).$$
   HINT: Consider $\{p\cdot\alpha, p\cdot\Sigma\}u_\pm(p)$, and you may appreciate the identity
$$\{\alpha^m, \epsilon_{ijk}\alpha^j\alpha^k\} = \begin{cases} 2\delta_{im}\epsilon_{ijk}\alpha^i\alpha^j\alpha^k, & i = m \\ 0, & i \neq m \end{cases}$$

8. Show that the expression for $\Delta T^{\alpha\beta}$ in Eq. (3.45) satisfies
$$\partial_\alpha \Delta T^{\alpha\beta} = 0.$$

4 Electromagnetic Interactions

4.1 Gauge Invariance

For complex fields, either bosonic or fermionic, an obvious symmetry involves rotating the field by a phase $e^{i\Lambda}$. Since the fields enter the Lagrangian as $\phi^* \cdots \phi$ or as $\bar{\psi} \cdots \psi$, nothing appears to be changed by such a phase change. However, if Λ depends on position and/or time this invariance disappears due to the derivatives acting on the phase, i.e.,

$$i\partial_\alpha(e^{i\Lambda}\psi) = e^{i\Lambda}(i\partial_\alpha - \partial_\alpha\Lambda)\psi. \tag{4.1}$$

Now, take it as a given that all derivatives acting on charged particles include the electromagnetic interaction,

$$i\partial_\alpha \rightarrow i\partial_\alpha - eA_\alpha, \tag{4.2}$$

where A is the electromagnetic field. In that case,

$$(i\partial_\alpha - eA_\alpha)(e^{i\Lambda}\psi) = e^{i\Lambda}(i\partial_\alpha - eA_\alpha - \partial_\alpha\Lambda)\psi. \tag{4.3}$$

Thus, the Lagrangian is invariant under transformations of the kind,

$$\psi \rightarrow e^{i\Lambda(x)}\psi, \;\; eA_\alpha \rightarrow eA_\alpha + \partial_\alpha\Lambda. \tag{4.4}$$

Actually, one needn't really add the extra term to A_α, because from Noether's theorem the conserved current is

$$j^\mu = i\frac{\partial\mathcal{L}}{\partial\partial_\mu\Lambda}, \tag{4.5}$$

which means that the change to $\mathcal{L}$ due to the transformation is

$$\Delta\mathcal{L} = j^\mu\partial_\mu\Lambda. \tag{4.6}$$

Since contributions to the action go as $i\int d^4x\mathcal{L}$, the contribution from $\Delta\mathcal{L}$ disappears as long as $\Lambda(x)$ is zero at some large distance, which can be seen by integrating by parts and assuming current conservation.

Thus, gauge invariance is more of a prescription for writing a theory, than some sort of derived quantity. It basically states that if the theory is to be invariant to gauge changes of the vector field, the vector field had better couple to conserved currents. Once you derive conserved currents from Noether's theorem, this becomes equivalent to stating that $i\partial_\alpha \rightarrow i\partial_\alpha - eA_\alpha$ when modifying a theory for electromagnetic interactions. Thus, for the Klein-Gordon equation, the Lagrangian is

$$\mathcal{L} = \phi^*(i\partial^\alpha - eA^\alpha)(i\partial_\alpha + eA_\alpha)\phi - m^2\phi^*\phi, \tag{4.7}$$

and the Dirac Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - e\not{A} - m)\psi. \tag{4.8}$$

EXAMPLE:
Solve for the ground state energies of a spinless bosonic charged particle in a constant magnetic field.

First, pick a gauge to describe a constant magnetic field in the z direction,

$$A_0 = A_x = A_z = 0,\ A_y = Bx.$$

Since the Lagrangian is invariant to translations along the y direction, solutions should be eigenstates of the operator P_y. Of course, solutions should also be eigenstates of P_z. Thus, for wave-function, we guess at the form,

$$\phi(x,y,z) = e^{-i\omega_t + ik_y y + ik_z z} e^{-(x-x_0)^2/2R^2}.$$

The guess at a Gaussian form comes because the Klein-Gordon equation has terms that go as x^2, which suggest a harmonic oscillator type solution.

Applying the K.G. equation,

$$\begin{aligned}\omega^2\phi &= k_z^2\phi + (k_y - eBx)^2\phi + m^2\phi + (-i\partial_x)^2\phi \\ &= (k_y^2 + k_z^2 + m^2)\phi + \left\{-2k_y eBx + e^2B^2x^2 - \frac{(x-x_0)^2}{R^4} + \frac{1}{R^2}\right\}\phi\end{aligned}$$

To be an eigenstate, all the x dependence must vanish. The two conditions to eliminate the terms proportional to x^2 and x are:

$$e^2B^2 = \frac{1}{R^4}, \quad -2k_y eB + 2\frac{x_0}{R^4} = 0.$$

Thus, x_0 and R are chosen so that

$$R^2 = \frac{1}{eB}, \quad x_0 = \frac{k_y}{eB}.$$

The K.G. equation is then

$$\omega^2\phi = \left(k_y^2 + k_z^2 + m^2 - \frac{x_0^2}{R^4} + \frac{1}{R^2}\right).$$

After substituting for x_0 and R, the k_y dependence also falls out, and

$$\omega^2 = k_z^2 + m^2 + eB.$$

Since all states with a given k_y have the same energy, it is interesting to ask how many degenerate states exist. The density of states is $L_y/(2\pi)$, so the number is for a fixed value of k_z is

$$N_{\text{degenerate}} = \frac{L_y}{2\pi}\int_{k_{y,\min}}^{k_{y,\max}}.$$

Since the center of the Gaussian wave function must be in a range, $0 < x_0 < L_x$, one can see that the range for k_y is $0 < k_y < eBL_x$, thus

$$N_{\text{degenerate}} = \frac{eBL_yL_x}{2\pi}.$$

This is a familiar result for Landau levels. One can also find another other sets of degenerate states by using the wave function for the excited states of a harmonic oscillator (For $n = 1$, $\phi \sim xe^{-x^2/2R^2}$). The result makes sense from the context of the uncertainty principle. For example, if you consider massless particles and $k_z = 0$, the uncertainty in the momentum is $\sqrt{eB}$, which equals $1/R$.

4.2 The Magnetic g-factor

The Dirac representation is the most reasonable for viewing properties of non-relativistic motion, such as the hydrogen atom, because the electron is largely non-relativistic, i.e., 13.6 eV $<<$ the electron's mass, 511 keV. Writing the spinor again as,

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{4.9}$$

the Dirac equation becomes

$$\begin{aligned} (i\partial_t - eA^0)\phi &= \sigma \cdot \pi\chi + m\phi \\ (i\partial_t - eA^0)\chi &= \sigma \cdot \pi\phi - m\chi, \\ \pi &\equiv -i\nabla - eA. \end{aligned} \tag{4.10}$$

Applying $(i\partial_t - eA^0 + m)$ to both sides of the upper equation allows one to substitute for χ using the lower equation, and obtain

$$(i\partial_t - eA_0)^2\phi = (\sigma \cdot \pi)^2\phi + m^2\phi. \tag{4.11}$$

Next, use the fact that

$$(\sigma \cdot \pi)^2 = \pi_i^2 + \frac{1}{4}[\sigma_i, \sigma_j][\pi_i, \pi_j]. \tag{4.12}$$

Commuting π_i and π_j gives $\epsilon_{ijk}eB_k$ and $[\sigma_i, \sigma_j] = 2\epsilon_{ijk}\sigma_k$, thus

$$(i\partial_t - eA_0)^2\phi = \pi^2\phi + m^2\phi - e\sigma \cdot B\phi. \tag{4.13}$$

As a next step, we will make a non-relativistic approximation. To do this, we define $\Phi \equiv \phi e^{imt}$. The oscillation of Φ will then be slow, i.e., $i\partial_t\Phi << m\Phi$. If one also tosses the terms or order $(eA_0)^2$, one finds

$$\begin{aligned} 2mi\partial_t\Phi &= \pi^2\Phi + 2meA_0\Phi - e\sigma \cdot B\Phi, \\ i\partial_t\Phi &= \frac{\pi^2}{2m}\Phi + eA_0\Phi - \frac{e}{2m}\sigma \cdot B\Phi. \end{aligned} \tag{4.14}$$

Normally, the coupling of the magnetic field to a current loop of angular momentum L from a single charge e would be (include $\hbar$ and c)

$$H_B = -\frac{e\hbar}{2mc}L \cdot B. \tag{4.15}$$

However, given that the spin operator is $S = \sigma/2$, one finds

$$H_B = -g\frac{e\hbar}{2mc}S \cdot B, \quad g = 2. \tag{4.16}$$

The fact that $g = 2$ is a non-trivial result that does much to validate the Dirac equation. From the context of the Schrödinger equation, the gyromagnetic factor is simply introduced as a free parameter with no explanation of why it would be near 2. In practice it differs from 2 slightly due to radiative corrections which can be calculated with higher-order perturbation theory.
If one would wish to stay away from the non-relativistic approximation, the coupling to magnetic fields can be derived more elegantly. Multiplying the Dirac equation by $(i\not{\partial} - e\not{A} + m)$ gives

$$\begin{aligned} (i\not{\partial} - e\not{A} + m)(i\not{\partial} - e\not{A} - m)\psi &= 0, \\ \left\{(i\partial - eA)^2 + \frac{1}{2i}\sigma^{\mu\nu}[i\partial_\mu - eA_\mu, i\partial_\nu - eA_\nu] - m^2\right\}\psi &= 0 \\ \left[(i\partial - eA)^2 - \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} - m^2\right]\psi &= 0. \end{aligned} \tag{4.17}$$

The term that provides the coupling of the spin to the electro-magnetic field is

$$-\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} = -g\frac{e}{2}(i\alpha\cdot E + \sigma\cdot B). \tag{4.18}$$

4.3 The Foldy-Wouthuysen Transformation and the Spin-Orbit Interaction

The terms of the previous section were of order $1/m$, the order $1/m^2$ and $1/m^3$ terms are especially important, such as the spin-orbit interaction. Calculating the next order is remarkably difficult. The difficulty is that the expansion is conceptually difficult, and once the method is understood, horribly messy.
The method for the solution is based on the Foldy-Wouthuysen transformation, which rotates the state by a unitary transformation,

$$\psi' = e^{iS}\psi, \tag{4.19}$$

so the that new Hamiltonian is

$$i\partial_t\psi' = H'\psi', \;\; H' = e^{iS}He^{-iS}. \tag{4.20}$$

The trick is to find the operator S so that H' has no components that mix upper and lower components in the Dirac representation, i.e., eliminate the terms proportional to the Dirac matrices α_i. For the case of a free particle, it is not particularly difficult. One simply guesses at the form,

$$S = \beta\alpha\cdot\hat{p}\theta, \tag{4.21}$$

and then solves for the angle θ. Note the β and α are the Dirac matrices. With this choice,

$$\begin{aligned} e^{iS} &= \cos\theta + (\beta\alpha\cdot\hat{p})\sin\theta, \\ H' &= (\cos\theta + \beta\alpha\cdot\hat{p}\sin\theta)(\alpha\cdot p + \beta m)(\cos\theta - \beta\alpha\cdot\hat{p}\sin\theta) \\ &= (\alpha\cdot p)\left(\cos 2\theta - \frac{m}{|p|}\sin 2\theta\right) + \beta(m\cos 2\theta + |p|\sin 2\theta). \end{aligned} \tag{4.22}$$

One then chooses $\tan 2\theta = |p|/m$ to eliminate the first term to get

$$H' = \beta\cos 2\theta(m + |p|^2/m) = \beta\sqrt{m^2 + |p|^2}. \tag{4.23}$$

The preceeding matrix β is necessary so that the negative eigenvalues also survive the procedure.
The situation become much messier once one includes interactions with the magnetic field and expresses the momentum operator as derivatives,

$$H = \alpha \cdot (p - eA) + \beta m + eA^0. \tag{4.24}$$

We define the even and odd terms (in α) as

$$\mathcal{O} \equiv \alpha \cdot (p - eA), \quad \mathcal{E} \equiv eA^0, \tag{4.25}$$

so that

$$H = \beta m + \mathcal{E} + \mathcal{O}. \tag{4.26}$$

Unfortunately, S cannot be obtained by guessing. Instead we must solve for S, and do so order-by-order in powers of $\mathcal{O}$ and $\mathcal{E}$, i.e, relative to the term βm in the Hamiltonian, this will turn out to be an expansion in $1/m$. Further, S will have an explicit time dependence, this means that

$$\begin{aligned} H\psi &= i\partial_t(e^{-iS}\psi') = e^{-iS}i\partial_t\psi' + (\partial_t e^{-iS})\psi', \\ i\partial_t\psi' &= \left[e^{iS}(H - i\partial_t)e^{-iS}\right]\psi', \\ H' &= \left[e^{iS}(H - i\partial_t)e^{-iS}\right]. \end{aligned} \tag{4.27}$$

Since neither H nor $\partial_t S$ commute with S, one must expand the expression for e^{iS} as a Taylor expansion in S, then commute the non-S terms outside so that $e^{iS}e^{-iS} = 1$ can be applied. This can be worked out with a moderate amount of pain to obtain,

$$\begin{aligned} H' &= H + i[S, H] - \frac{i^2}{2}[S, [S, H]] + \cdots \frac{i^n}{n!}[S, [S, \cdots [S, H] \cdots]] \\ &\quad -\dot{S} - \frac{i}{2}[S, \dot{S}] + \cdots \frac{i^n}{n!}[S, S, \cdots [S, H] \cdots]]. \end{aligned} \tag{4.28}$$

To first order in S,

$$H' \approx \beta m + \mathcal{E} + \mathcal{O} + i[S, \beta]m. \tag{4.29}$$

To eliminate the odd terms, one needs to choose S so that $i[S, \beta] = -\mathcal{O}$. A little bit of thinking convinces one that

$$S = \frac{-i}{2m}\beta\mathcal{O}. \tag{4.30}$$

Now one can rewrite H' including all the double commutators that were neglected before. These bring in some odd terms in $\mathcal{O}$, but they are of order $\mathcal{O}^3$. One can now repeat the procedure, treating H' as if it were the Hamiltonian, and re-defining $\mathcal{E}$ and $\mathcal{O}$ to represent the even/odd terms in H'. This procedure will result in a new Hamiltonian, H', and can then be repeated again to find the Hamiltonian H''' which will be relativistic to order $1/m^3$. The result, after a tremendous amount of pain, is:

$$\begin{aligned} H''' &= \beta\left(m + \frac{(p - eA)^2}{2m} - \frac{p^4}{8m^3}\right) + eA^0 - \frac{e}{2m}\beta\Sigma \cdot B \\ &\quad -\frac{ie}{8m^2}\Sigma \cdot (\nabla \times E) - \frac{e}{4m^2}\Sigma \cdot E \times p - \frac{e}{8m^2}(\nabla \cdot E), \end{aligned} \tag{4.31}$$

where $\Sigma_i \equiv (-i/2)\epsilon_{ijk}\gamma_j\gamma_k$. Now, to discussed the terms individually:
(1) $p^4/8m^3$ is simply the order $1/m^3$ term in the expansion of $\sqrt{m^2+p^2}$.
(2) $(e/2m)\Sigma\cdot B$ is the term that couples the spin, Σ, to the magnetic field discuss in the previous section.
(3) $(-ie/8m^2)\Sigma\cdot(\nabla\times E) - (e/4m^2)\Sigma\cdot E\times p$ is the spin-orbit energy. For the hydrogen atom, or any spherhically symmetric potential, the first term is zero and the second term can be simplified by seeing

$$E = E_r r/r,\ \ \Sigma\cdot E\times p = -\frac{E_r}{r}\Sigma\cdot(r\times p) = -\frac{E_r}{r}\Sigma\cdot(L). \tag{4.32}$$

which then becomes the spin-orbit interaction,

$$H_{\text{s.o.}} = \frac{e}{4m^2}\frac{E_r}{r}\Sigma\cdot L. \tag{4.33}$$

The spin-orbit term is often motivated within the context of a Schrödinger equation by considering the magnetic field in the frame of the electron, which is odd given that the discussion concerns wavefunctions not classical point particles. In that frame, it looks like the spin-orbit interaction. These kinds of hand-waving arguments are more a matter of desperation than insight.
(4) The last term, $(e/8m^2)(\nabla\cdot E)$, is particularly funky. It is known as the Darwin term. From Maxwell's equations, it looks as if it couples to the charge density. Again, there are some hand-waving arguments to physically motivate the term. One could say that for a point particle, stating that the particle is exactly at position r has some uncertainty determined by the Compton wavelength, $\lambda\sim 1/m$. Thus, the Coulomb energy is a little uncertain. Again this is simply another desperate attempt to understand something that is really simply a matter of "It fell out of the math".
One interesting side note regarding the spin-orbit interaction is that if you had a scalar potential, rather than the zero$^{\text{th}}$ component of a vector potential, the sign would switch, i.e.,

$$\partial_r eA_0 \to -\partial_r M(r). \tag{4.34}$$

This becomes important for nuclear physics because the nuclear force has both vector and scalar components. The scalar part is attractive and comes mainly from two-pion exchange while the vector part is associated with omega exchange and is repulsive. Both have strengths on the order of ∼300 MeV, but they largely cancel to result in a typical potential well depth of ~ -35 MeV. Although the potentials largely cancel, the spin-orbit contributions are of the same sign. Thus compared to using simply a vector, or simply a scalar potential, the spin-orbit contribution is 20 times larger than expected. Nuclear shell structure was a mystery until Maria Goeppert Mayer ran her nuclear shell model with extremely large spin-orbit couplings just to see what happened. Lo and behold, she explained the nuclear magic numbers.

4.4 The Hydrogen Atom

If you don't look too closely, the energy levels of the hydrogen atom depend only on the principal quantum number n,

$$E_n = \frac{-13.6}{n^2}\ \text{eV}. \tag{4.35}$$

The degeneracies are such that all values of ℓ are allowed that are less than n. For the $n = 2$ level, one would have both $\ell = 0$ (s-wave) and $\ell = 1$ (p-wave). Further, the p-wave solutions could have $j = 1/2$ or $j = 3/2$ solutions. The spin-orbit interaction is proportional to $\boldsymbol{L} \cdot \boldsymbol{S}$, and since

$$-\boldsymbol{L} \cdot \boldsymbol{S} = \frac{\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2}{2} = -\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{2}, \tag{4.36}$$

the interaction breaks the degeneracy between the two p-wave solutions with different j. If one averages over the $j = 1/2$ and $j = 3/2$ solutions (weighing by $(2j + 1)$), one can see that the levels are split by the spin-orbit interaction, but the average energy stays in place. Thus, one would expect the average energies of the 2p solutions to remain equal to that of the 2s state. The splitting is referred to as *fine structure*.

However the energies of the 2s states differ slightly from that of the 2p states. This comes from perturbative processes involving a photon being emitted and reabsorbed by the electron. The shift can be calculated to high precision and is known as the *Lamb shift*. It is proportional to e^2 (as compared to the spin-orbit interaction which is proportional to e) and has about 10% of the strength of the spin-orbit piece.

Finally, there are also perturbative processes involving the proton emitting a real photon (not the classical field) and having it reabsorbed by the electron. This turns out to be a spin-spin interaction, involving the spins of both the electron and the proton. Since it behaves as $\boldsymbol{S_p} \cdot \boldsymbol{S_e}$, it also ends up depending on $\boldsymbol{J}$, but in this case it is the $\boldsymbol{J}$ including the coupling to the protons spin. Even for the 1s solution, there are $\boldsymbol{J} = 0$ and $\boldsymbol{J} = 1$ solutions. Since the proton is heavier, this term is smaller than that of the Lamb shift roughly by the ratio of the proton and electron mass. The $\boldsymbol{J} = 0$ and $\boldsymbol{J} = 1$ levels of the 1s levels in hydrogen are split by $\sim$ one thousandth as much as the Lamb shift. This splitting is known as *hyperfine structure*.

4.5 Homework Problems

1. Consider the massless Dirac equation in a constant magnetic field described by the vector potential, $\boldsymbol{A_x} = \boldsymbol{A_z} = \boldsymbol{A_0} = 0$, $\boldsymbol{A_y} = \boldsymbol{Bx}$. For the questions below, use the chiral representation of the Dirac equation, which allows you to use two-by-two matrices when there is no mass. In that representation,

$$\vec{\boldsymbol{\alpha}} = \begin{pmatrix} \vec{\boldsymbol{\sigma}} & 0 \\ 0 & -\vec{\boldsymbol{\sigma}} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}.$$

   Because there is no mass, there is no mixing of upper and lower components and we can simply consider two-component spinors, and for the upper components, we can simplify the Dirac matrices, $\vec{\boldsymbol{\alpha}} = \vec{\boldsymbol{\sigma}}$.

   (a) Write the Hamiltonian in terms of $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{\partial_i}$, $\boldsymbol{B}$, and the $\boldsymbol{\sigma}$ matrices.

   (b) Try a solution of the form,

$$\boldsymbol{\psi} = \boldsymbol{e}^{-\boldsymbol{iEt}+\boldsymbol{ik_z z}+\boldsymbol{ik_y y}} \boldsymbol{e}^{-(\boldsymbol{x}-\boldsymbol{x_0})^2/2\boldsymbol{R}^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

   and solve for the eigen-energy.

2. Consider the Klein-Gordon equation for a charged scalar field interacting with an electromagnetic field.

   (a) Derive the stress-energy tensor beginning with the prescription in the first line of Eq. (3.44).

   (b) The expression in (a) is not symmetric. Show that one can add a term

   $$\boldsymbol{\Delta T^{\alpha\beta} = -e j^{\alpha} A^{\beta}}$$

   to the stress-energy tensor to obtain a symmetric expression.

   (c) Show that

   $$\boldsymbol{\partial_{\alpha} \Delta T^{\alpha\beta} = 0.}$$

5 Field Theories

5.1 Non-Relativistic Field Theory

First, we consider the non-relativistic case. In that case we consider creation and destruction operators, which are based on raising and lowering operators used to address the harmonic oscillator problem. In that case, one can consider N harmonic oscillators, with raising and lowering operators $c_i^\dagger$ and c_i. Here, i runs from i to N, and the commutation relations are:

$$[c_i, c_j^\dagger] = \delta_{ij}. \tag{5.1}$$

To see the correspondence with field operators, consider operators that "raise" or "lower" the number of particles at point r. Let them commute similarly as above,

$$[\Psi(r), \Psi^\dagger(r')] = \delta^3(r - r'). \tag{5.2}$$

Further, one can consider the Fourier transform of $\Psi(r)$,

$$a_k = \frac{1}{\sqrt{\Omega}} \int d^3r e^{-ik\cdot r} \Psi(r), \tag{5.3}$$

where we are considering a volume Ω and the values k correspond to eigenstates in an infinite well of that volume, i.e., $k_x L_x = n_x \pi,\ k_y L_y \cdots$. One can then calculate the commutator for a_k,

$$\begin{aligned}
[a_k, a_{k'}^\dagger] &= \frac{1}{\Omega} \int d^3r d^3r'\ e^{-ik\cdot r + ik'\cdot r'} [\Psi(r), \Psi^\dagger(r')] \\
&= \frac{1}{\Omega} \int d^3r d^3r'\ e^{-ik\cdot r + ik'\cdot r'} \delta^3(r - r') \\
&= \frac{1}{\Omega} \int d^3r\ e^{-i(k-k')\cdot r} \\
&= \delta_{k,k'}.
\end{aligned} \tag{5.4}$$

Thus, every momentum mode, or equivalently every point in coordinate space, is treated as if it had a separate harmonic-oscillator basis. Similarly, one can invert Eq. (5.3) and find $\Psi(r)$ in terms of $a(k)$,

$$\begin{aligned}
\frac{1}{\sqrt{\Omega}} \sum_k e^{-ik\cdot r} a_k &= \sum_k \frac{1}{\Omega} \int d^3r\ e^{-ik\cdot(r-r')} \Psi(r') &(5.5)\\
&= \Psi(r). &(5.6)
\end{aligned}$$

The Hamiltonian for free particles can be found by summing over the number of particles for each momentum mode multiplied by the energy of particles in that mode,

$$H = \sum_k \frac{k^2}{2m} a_k^\dagger a_k. \tag{5.7}$$

Inserting the expression for a_k in Eq. (5.3),

$$\begin{aligned} H &= \frac{1}{\Omega}\sum_k \int d^3r d^3r'\, e^{-ik\cdot(r-r')}\Psi^\dagger(r)\frac{k^2}{2m}\Psi(r') \\ &= \frac{1}{\Omega}\sum_k \int d^3r d^3r'\, e^{-ik\cdot(r-r')}\Psi^\dagger(r)\frac{-\nabla^2}{2m}\Psi(r') \\ &= \int d^3r \Psi^\dagger(r)\frac{-\nabla^2}{2m}\Psi(r), \end{aligned} \tag{5.8}$$

where the second step involved integrating by parts and the third step invoked the inverse Fourier transform, $\sum_k e^{-ik\cdot(r-r')} = \Omega\delta^3(r-r')$.
Equation (5.7) looks like the usual expression for the energy of a free particle in the Schrödinger equation, except in this case Ψ and $\Psi^\dagger$ are operators, NOT wavefunctions. To calculate the energy of a system in state $|\phi\rangle$, one would need to calculate

$$E = \langle\phi|H|\phi\rangle. \tag{5.9}$$

The energy E would include the energy of all the particles involved in the system, whereas the Schrödinger equation is limited to discussions of a single particle. Thus, basically all many-body problems are formulated in terms of field theory, and since relativistic physics always includes processes that create particle-antiparticle pairs, they too are expressed in the language of field theory.
The expressions above were calculated assuming a finite volume Ω. The more usual formalism used is to let $\Omega \to \infty$ and to use operators $a(k)$ and $a^\dagger(k)$ that commute to give a Dirac delta, rather than a Kronecker delta,

$$[a(k), a^\dagger(k')] = (2\pi)^3\delta^3(k-k'), \tag{5.10}$$

where the momenta are then considered to be continuous. Note that k appears as a subscript for the discrete case, a_k, whereas in the continuous limit k appears as $a(k)$. The Fourier transforms of the field operators in the continuum limit are

$$\begin{aligned} a(k) &= \int d^3r\, e^{-ik\cdot r}\Psi(r), \\ \Psi(r) &= \frac{1}{(2\pi)^3}\int d^3k e^{ik\cdot r}a(k), \end{aligned} \tag{5.11}$$

and the Hamiltonian in momentum space is:

$$H = \int \frac{d^3k}{(2\pi)^3}\frac{k^2}{2m}a^\dagger(k)a(k). \tag{5.12}$$

The Hamiltonian written in terms of coordinate space operators looks exactly the same as before.

EXAMPLE:
Consider the state

$$|\phi\rangle = \int \frac{d^3r}{(2\pi r^2)^{3/4}}\, e^{-r^2/4R^2}\Psi^\dagger(r)|0\rangle.$$

Using the standard definitions,

$$|r\rangle = \Psi^\dagger(r)|0\rangle, \quad |k\rangle = a^\dagger(k)|0\rangle,$$

a) Calculate $\langle r|\phi\rangle$ and $\int d^3r\ |\langle r|\phi\rangle|^2$.

$$\begin{aligned}
\langle r|\phi\rangle &= \langle 0|\Psi(r)\int \frac{d^3r'}{(2\pi R^2)^{3/4}} e^{-r'^2/4R^2}\Psi^\dagger(r')|0\rangle \\
&= \frac{1}{(2\pi R^2)^{3/4}} e^{-r^2/4R^2} \\
\int d^3r\ |\langle r|\phi\rangle|^2 &= \int d^3r\ \frac{1}{(2\pi R^2)^{3/4}} e^{-r^2/2R^2} \\
&= 1.
\end{aligned}$$

b) Calculate $\langle k|\phi\rangle$ and $\int (d^3k/(2\pi)^3)\ |\langle k|\phi\rangle|^2$.

$$\begin{aligned}
\langle k|\phi\rangle &= \langle 0|\int d^3r e^{-ik\cdot r}\Psi^\dagger(r)\int \frac{d^3r'}{(2\pi R^2)^{3/4}} e^{-r'^2/4R^2}\Psi^\dagger(r')|0\rangle \\
&= \int \frac{d^3r}{(2\pi R^2)^{3/4}}\ e^{-ik\cdot r} e^{-r^2/4R^2} \\
&= e^{-k^2R^2}\int \frac{d^3r}{(2\pi R^2)^{3/4}}\ e^{-(r+2ikR^2)^2/4R^2} \\
&= (8\pi R^2)^{3/4} e^{-k^2R^2} \\
\int \frac{d^3k}{(2\pi)^3}\ |\langle k|\phi\rangle|^2 &= \frac{1}{[2\pi(4/R^2)]^{3/2}}\int d^3k\ e^{-k^2/[2(R^2/4)]} \\
&= 1.
\end{aligned}$$

5.2 Relativistic Scalar Fields

For a complex relativistic scalar field, the current density was $\mathcal{R}e\phi^*(i\partial^\mu)\phi$. Thus ϕ should have dimensions of inverse mass. Guessing at an expression for the field operator Φ,

$$\Phi(x) = \int \frac{d^3k}{2E_k(2\pi)^3}\left(a(k)e^{-ik\cdot x} + b^\dagger(k)e^{ik\cdot x}\right). \tag{5.13}$$

The choice of putting in the $1/2E_k$ is because $d^3k/2E_k$ is Lorentz invariant. To see this,

$$\int d^4k\ \delta(k^2 - m^2) = \int d^3k \frac{1}{\left.\frac{\partial k^2}{\partial k_0}\right|_{k_0=E_k}} = \int \frac{d^3k}{2E_k}. \tag{5.14}$$

The choice of assigning a creation operator $b^\dagger$ rather than b is because the operators b correspond to negative energy particles, i.e., the negative energy will be related to moving a particle. In a

moment we will also see that these particles have opposite sign, i.e., they are anti-particles. We now take a stab at writing down invariant commutation relations,

$$[a(k), a^\dagger(k')] = (2\pi)^3 2E_k \delta^3(k-k'), \tag{5.15}$$

with the same rules for b with $b^\dagger$ and with all other commutators equal to zero (such as between a and a, between $a^\dagger$ and $a^\dagger$, between a and b, between a and $b^\dagger$, etc. One can see that these are invariant because

$$\int \frac{d^3k}{(2\pi)^3 2E_k} [a(k), a^\dagger(k')] = 1, \tag{5.16}$$

irrespective of the reference frame. Continuing with the guess, let's calculate the integrated current at at time $x_0 = 0$,

$$\begin{aligned} Q^\mu &= \int d^3x\, j^\mu(x) = \int d^3x\, \left\{\Phi^\dagger(x)\left(i\partial^\mu\Phi(x)\right) + \left(-i\partial^\mu\Phi^\dagger(x)\right)\Phi(x)\right\} \\ &= \mathcal{R}e \int d^3x \frac{d^3k}{(2\pi)^3 2E_k}\frac{d^3k'}{(2\pi)^3 2E_{k'}}(k'^\mu + k^\mu) e^{i(k-k')\cdot x}[a^\dagger(k') + b(-k')][a(k') - b^\dagger(-k')] \\ &= \mathcal{R}e \int \frac{d^3k}{(2\pi)^3 2E_k}\frac{k^\mu}{E_k}\left\{a^\dagger(k)a(k) - b(-k)b^\dagger(-k) - e^{2iE_kt}a^\dagger(k)b^\dagger(-k) + e^{-2iE_kt}b(-k)a(k)\right\} \end{aligned} \tag{5.17}$$

Taking the real part eliminates the last two terms, and since one integrates over k, the $-k$s can be switched to $+k$ in the terms with bs and $b^\dagger$s. This yields,

$$Q^\mu = \int \frac{d^3k}{(2\pi)^3 2E_k}\frac{k^\mu}{E_k}\left[a^\dagger(k)a(k) - b(k)b^\dagger(k)\right]. \tag{5.18}$$

The $a^\dagger(k)a(k)$ term is simply the density of positive particles per invariant momentum space, $dN/(d^3k/E_k)$. The second term has the creation operators backwards, but can be understood by the commutation relation,

$$b(k)b^\dagger(k) = b^\dagger(k)b(k) + (2\pi)^3 2E_k\delta^3(k-k). \tag{5.19}$$

Now, the integrated current Q^μ looks like the integrated number of particles (the $a^\dagger a$ term) minus the number current for anti-particles (the $b^\dagger b$ term), plus an extra term. The extra term would even contribute to the vacuum. Further, the contribution is infinite. To see this, we first use the relation $\int d^3x e^{ik\cdot x} = (2\pi)^3\delta(k)$ in the limit of infinite volume to replace $\delta^3(0)$ with $\Omega/(2\pi)^3$. This gives the contribution of the final term in the vacuum as:

$$\langle 0|Q^0|0\rangle = -\Omega\int\frac{d^3k}{(2\pi)^3}. \tag{5.20}$$

It appears that for every discrete momentum mode k, there is a charge of -1 lying in the vacuum, and since there are an infinite number of such modes, there is an infinite charge. Further, one could repeat the operations above for the integrated components of the stress energy tensor. This could include the energy density, $H = \int d^3x\, T^{00}$, where T^{00} was derived earlier in terms of wave functions. If one did this, the integrated energy of the vacuum would come out as

$$\langle 0|H|0\rangle = \langle 0|T^{00}|0\rangle = \Omega\int\frac{d^3k}{(2\pi)^3}E_k. \tag{5.21}$$

This can be interpreted by considering each mode as a harmonic oscillator with two modes (real and complex). Each contributes $(1/2)\hbar\omega$, or in this case $E_k/2$ to the vacuum energy. Note that there is no net momentum in the vacuum because the modes k and $-k$ cancel one another.
Both the infinite vacuum energy and the infinite charge are disturbing. Subtracting away the infinite vacuum energy is distressing because it would contribute to the energy density in gravitational theories. The infinite charge seems even more bizarre. For one thing, the charge and energy are four-vectors. If there were an infinite charge density, wouldn't there be an infinite current unless one were in a very specific frame? Such sicknesses are common in field theory. As long as the sickness can be isolated by adding an arbitrary term to cancel them, all is well. In this case one can imagine an infinite background charge of the opposite sign, which just magically cancels the contribution above. Even more complicated infinities appear when one begins to account for screening of interactions. Handling these infinities is the principal goal of renormalization theory. It is also worth noting that in fermionic field theories, the infinities described above will be equal and opposite. Thus, one of the motivations of super-symmetry, which postulates equal numbers of bosons and fermions of the same charges, is to cancel such infinities.

5.3 Real Scalar Fields

The only "fundamental" scalar field in the standard model (at least before symmetry breaking) is the Higgs boson. However, the Higgs is neutral. To account for the fact that there are not two charges, one can simply use the field operators as the previous section, but with the condition

$$b(k) \to a(k), \;\; b^\dagger(k) \to a^\dagger(k). \tag{5.22}$$

The field operators are then,

$$\Phi(x) = \Phi^\dagger(x) = \int \frac{d^3k}{(2\pi)^3 2E_k} \left[e^{-ik\cdot x}a(k) + e^{ik\cdot x}a^\dagger(k)\right]. \tag{5.23}$$

Note that there is no current associated with such fields. If you proceeded with the current operator as was done in the previous section, one would find terms of the form $(a^\dagger(k)a(k) - a^\dagger(k)a(k))$, rather than terms like $(a^\dagger(k)a(k) - b^\dagger(k)b(k))$.

5.4 Real Vector Fields

The most common bosonic field to be considered is a vector field, the photon. For any given momentum mode, free photons have two polarizations. For massive particles, one expects three polarizations. This all becomes rather difficult to express in a manifestly covariant way, and as will be seen later on, the confusion is tangled up with issues of gauge invariance. To get started, we will first consider the massless (photon) case. Again, beginning with a guess for the form of the field operators,

$$A^\mu(x) = \int \frac{d^3k}{(2\pi)^3 2E_k} \left[e^{-ik\cdot x}a^{(\mu)}(k) + e^{ik\cdot x}a^{(\mu)\dagger}(k)\right]. \tag{5.24}$$

The creation and destruction operators then carry the index describing the vector nature of the fields. Since the operators have the extra indices, one must write down the commutation relations in a gauge-invariant way. One common way is (the Feynmann gauge):

$$[a^\mu(k), a^{\nu\dagger}(k')] = -g^{\mu\nu}2E_k(2\pi)^3\delta^3(k-k'). \tag{5.25}$$

These are bizarre, because it appears that there are four free fields, instead of two, and that one of them has commutation law with the wrong sign. For a commutation law with the wrong sign, one ends up with a negative-norm state for $\mu = 0$. In the limit of discrete modes,

$$\langle 0|a_k^0 a_k^{0\dagger}|0\rangle = -1. \tag{5.26}$$

The negative norm state complicates all our normal ideas about Hermitian Hamiltonians and unitarity. In particular, of all the principles and symmetries of quantum mechanics, the only absolute principle that should be obtained is unitarity, i.e., that the net probability summed over all states is unity – and stays that way. The solution to the puzzle comes from simultaneously ignoring one of the positive-norm helicities. Of the three positive-norm helicities, two are transverse and one is longitudinal. For a given momentum state k, pointing along the z axis, the normalizations are:

$$\epsilon^{(0)}(k) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ \epsilon^{(1)}(k) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ \epsilon^{(2)}(k) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ \epsilon^{(3)}(k) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{5.27}$$

With these definitions, one can rewrite the field operators in Eq. (5.24) as:

$$A^\mu(x) = \sum_{(\lambda)} \int \frac{d^3k}{(2\pi)^3 2E_k} \left[e^{-ik\cdot x} \epsilon^\mu_{(\lambda)}(k) a^{(\lambda)}(k) + e^{ik\cdot x} \epsilon^{*\mu}_{(\lambda)}(k) a^{(\lambda)\dagger}(k) \right]. \tag{5.28}$$

However, there is no particular reason that the sum over polarizations needs to be expressed in the same basis for every k. Instead $\lambda = 1, 2$ could refer to any two unit vectors that are orthogonal to k, while $\lambda = 3$ could refer to the polarization parallel to k. The polarization (0) along with the polarization parallel to k are referred to as the longitudinal polarizations. Proving unitarity involves showing that $\sum_{\psi_f} |\langle \psi_f | e^{-iHt} | \psi_0 \rangle|^2 = 1$. One can write the states as a product of longitudinal and transverse states, $|\psi\rangle = |\psi_T\rangle|\psi_L\rangle$. When one calculates a probability, one can ask the question, "Summing over all states ψ_L, What is the probability of going into the state ψ_T?" Whereas one might find the possibility that the probability that going into a specific longitudinal state has negative probability (and can depend on the choice of gauge), summing over all longitudinal states always yields a positive probability for any state $|\psi_T\rangle$ (and is independent of gauge). We will return to this issue later.

One can also consider massive vector particles. Beginning with the Lagrangian,

$$L = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} m^2 A^2, \ \ F^{\mu\nu} = \partial^\mu A_\nu - \partial_\nu A^\mu, \tag{5.29}$$

the equations of motion become

$$-\partial_\rho F^{\rho\mu} + m^2 A^\mu = 0. \tag{5.30}$$

Taking the divergence, and using $\partial_\rho \partial_\mu F^{\rho\mu} = 0$ due to F being anti-symmetric, one finds

$$\partial \cdot A = 0. \tag{5.31}$$

Now, using the identity $\partial_\rho F^{\rho\mu} + \partial_\mu \partial \cdot A = \partial^2 A$, the equations of motion can be written as:

$$-\partial^2 A^\mu = m^2 A^\mu, \ \ \partial \cdot A = 0. \tag{5.32}$$

This would be satisfied by writing the field operator with only three polarizations, and with $\epsilon^{(0)}$ going to zero in the frame where $k = 0$. The massive case no longer respects gauge invariance.

5.5 Coherent States, Path Integrals and Lagrangians

Thus far, we have not made a connection between a Lagrangian and expressing the fields in terms of creation and destruction operators. To do so, we first consider *coherent states*. A coherent state is state defined not by a particle number, but rather by being an eigenstate of the destruction operator. For a single mode, the state is defined by a complex number η,

$$|\eta\rangle = \exp\{\eta a^\dagger - \eta^* a\}|0\rangle. \tag{5.33}$$

Since the argument of the exponential is anti-Hermitian, the state $|\eta\rangle$ is normalized to unity, $\langle\eta|\eta\rangle = 1$. The Baker-Campbell-Hausdorff lemma says that for two operators A and B that commute to a number, one can say $e^{A+B} = e^A e^B e^{-[A,B]/2}$, which allows one to restate $|\eta\rangle$ as

$$|\eta\rangle = e^{-|\eta|^2/2} e^{\eta a^\dagger} e^{-\eta^* a}|0\rangle = |\eta\rangle = e^{-|\eta|^2/2} e^{\eta a^\dagger}|0\rangle. \tag{5.34}$$

The factor $e^{-|\eta|^2/2}$ is a normalization factor. The unsettling aspect of coherent states is that they are not eigenstates of the number operator, $a^\dagger a$. However they are eigenstates of the destruction operator,

$$\begin{aligned} a|\eta\rangle &= e^{-|\eta|^2/2} a \left\{1 + (\eta a^\dagger) + \frac{(\eta a^\dagger)^2}{2} + \frac{(\eta a^\dagger)^3}{3!} + \cdots\right\}|0\rangle \\ &= e^{-|\eta|^2/2} \eta \left\{1 + (\eta a^\dagger) + \frac{(\eta a^\dagger)^2}{2} + \frac{(\eta a^\dagger)^3}{3!} + \cdots\right\}|0\rangle \\ &= \eta|\eta\rangle. \end{aligned} \tag{5.35}$$

EXAMPLE:
Calculate the the overlap $\langle\eta'|\eta\rangle$.

Expanding the bra and ket,

$$\begin{aligned} \langle\eta'|\eta\rangle &= e^{-|\eta'|^2/2 - |\eta|^2/2} \langle 0| e^{\eta'^* a} e^{\eta a^\dagger}|0\rangle \\ &= e^{-|\eta'|^2/2 - |\eta|^2/2} \langle 0| e^{\eta'^* \eta} e^{\eta a^\dagger}|0\rangle \\ &= e^{-|\eta'|^2/2 - |\eta|^2/2} e^{\eta'^* \eta}. \end{aligned}$$

Defining $\eta' = \eta + \delta\eta$, one can write the overlap as

$$\langle\eta'|\eta\rangle = e^{-|\delta\eta|^2/2} e^{(\delta\eta^* \eta - \eta^* \delta\eta)/2}. \tag{5.36}$$

Since the latter exponential is a phase, one can see that the magnitude of the overlap is maximized, and equal to unity, for the case that $\eta = \eta'$.

EXAMPLE:
Calculate the matrix element $\langle m|a^\dagger a^3|\eta\rangle$.

First use the fact that a coherent $|\eta\rangle$ is an eigenstate of a,

$$a^3|\eta\rangle = \eta^3|\eta\rangle.$$

Next, use the fact that for an eigenstate of particle number,

$$\langle m|a^\dagger = \sqrt{m}\langle m-1|.$$

Expanding the coherent state in a Taylor expansion,

$$|\eta\rangle = e^{-|\eta|^2/2}\left\{\sum_{n=0}^{\infty}\frac{(\eta a^\dagger)^n}{n!}\right\}|0\rangle = e^{-|\eta|^2/2}\left\{\sum_{n=0}^{\infty}\frac{\eta^n}{\sqrt{n!}}|n\rangle\right\}.$$

Finally, taking the overlap,

$$\langle m|a^\dagger a^3|\eta\rangle = \sqrt{m}\frac{\eta^{m+2}}{\sqrt{(m-1)!}}e^{-|\eta|^2/2}.$$

For eigenstates of the number operator, completeness is stated as $\sum_n \langle m|n\rangle\langle n|m'\rangle = \delta_{mm'}$. Even though the states are not orthogonal (see example above), there is a completeness relation,

$$\int \frac{d\eta_r d\eta_i}{\pi}\,\langle m|\eta\rangle\langle\eta|m'\rangle = \delta_{mm'}, \tag{5.37}$$

where η_r and η_i are the real and imaginary parts of η. To prove Eq. (5.37), one can expand the exponentials used to define the coherent state,

$$\int \frac{d\eta_r d\eta_i}{\pi}\,\langle m|\eta\rangle\langle\eta|m'\rangle = \int \frac{d\eta_r d\eta_i}{\pi}\,\langle 0|\frac{\eta^m}{\sqrt{m!}}|\eta\rangle\langle\eta|\frac{\eta^{*m'}}{\sqrt{m'!}}|0\rangle \tag{5.38}$$

$$= \int \frac{d\eta_r d\eta_i}{\pi}\,\frac{\eta^{*m'}\eta^m}{\sqrt{m!m'!}}e^{-|\eta|^2} \tag{5.39}$$

Here $\eta = |\eta|e^{i\phi}$. Using the fact that $\int d\phi e^{i(n-n')\phi} = 2\pi\delta_{nn'}$,

$$\int \frac{d\eta_r d\eta_i}{\pi}\langle m|\eta\rangle\langle\eta|m'\rangle = \int 2|\eta|d|\eta|e^{-|\eta|^2}\sum_m \frac{|\eta|^{2m}}{m!}\delta_{mm'} \tag{5.40}$$

$$= \delta_{mm'}.$$

Using the basis of coherent states, the evolution operator, e^{-iHt}, can be considered as a sum over paths, $\eta(t)$, through the basis. To see this, we breakup the time t into N segments of length $\delta t = t/N$, and take the limit of $N \to \infty$.

$$\langle\alpha|e^{-iHt}|\beta\rangle = \langle\alpha|e^{-iH\delta t}|\eta_1\rangle\int\frac{d\eta_{1r}d\eta_{1i}}{\pi}\langle\eta_1|e^{-iH\delta t}|\eta_2\rangle\int\frac{d\eta_{2r}d\eta_{2i}}{\pi}\langle\eta_2|\cdots|\beta\rangle \tag{5.41}$$

$$\begin{aligned}\langle\eta+\delta\eta|e^{-iH\delta t}|\eta\rangle &= e^{-iH(a^\dagger\to\eta^*,a\to\eta)\delta t}\langle\eta+\delta\eta|\eta\rangle \\ &= e^{(\eta^*\delta\eta-\delta\eta^*\eta)/2}e^{-iH(a^\dagger\to\eta^*,a\to\eta)\delta t} \\ &= e^{-i[i\eta^*\dot\eta/2-i\dot\eta^*\eta/2+H]\delta t}.\end{aligned}$$

The steps above used the fact that $\delta\eta$ and δt were small as $N \to \infty$. The expression for the overlap of two coherent states, $\langle\eta'|\eta\rangle$ derived in an earlier example, was also employed and take in the limit that $\delta\eta$ was small. Rather than integrating over η_r and η_i, we can define the real variables,

$$p = -i(\eta - \eta^*)/\sqrt{2}, \;\; q = (\eta + \eta^*)/\sqrt{2}, \;\; \eta = (q + ip)/\sqrt{2}, \;\; \eta^* = (q - ip)/\sqrt{2}. \quad (5.42)$$

With these variables, the path integrals become,

$$\begin{aligned} \langle\alpha|e^{-iHt}|\beta\rangle &= \langle\alpha|e^{-iH\delta t}|p_1, q_1\rangle \int \frac{dp_1 dq_1}{2\pi} \langle p_1, q_1| \\ &\quad \cdot e^{-iH\delta t}|p_2, q_2\rangle \int \frac{dp_2 dq_2}{2\pi} \langle p_2, q_2| \cdots |\beta\rangle \\ \langle p + \delta p, q + \delta q|e^{-iH\delta t}|p, q\rangle &= e^{i[(p\dot{q} - q\dot{p})/2 - H]\delta t}. \end{aligned} \quad (5.43)$$

This can now be recognized as a path integral,

$$\langle\alpha|e^{-iH\delta t}|\beta\rangle = \int \left(\Pi_{i=1,N} \frac{dp_i dq_i}{2\pi}\right) \exp\left\{i \int dt \, L(p(t), q(t))\right\} \quad (5.44)$$

$L(p, q) = (p\dot{q})/2 - (q\dot{p})/2 - H(p, q).$

The function $L(p, q)$ can equivalently be written

$$L(p, q) = p\dot{q} - H(p, q), \quad (5.45)$$

because the difference in the two expressions for L can be written as $(d/dt)(pq/2)$, which when integrated over time leads to a path-independent phase. In this form one can identify L as the Lagrangian.

If the path integral above were to be multidimensional, the derivation would proceed identically, only with significantly more pain. The variables p and q would go to p_ℓ, q_ℓ and the integrals over paths would become

$$\int \left(\Pi_{i=1,N} \frac{dp_i dq_i}{2\pi}\right) \to \int \left(\Pi_{i=1,N} \Pi_{\ell_i} \frac{dp_{\ell_i} dq_{\ell_i}}{2\pi}\right). \quad (5.46)$$

In a field theory, one assigns a field q_ℓ to each point in coordinate space. Of course, there are an infinity of such points, which may seem to make the expression impractical. However, in lattice gauge theory space and time are actually discretized so that such integrals can be performed via Markov-chain Monte Carlo (though usually done for thermal calculations where time goes along the imaginary axis, i.e., $e^{-iHt} \to e^{-\beta H}$.

5.6 Canonical Quantization

We are now in position to state how the quantized field operators can be extracted from the Lagrangian. If the Lagrangian is written in terms of q and $\dot{q}$, one can equivalently write it as

$$L(q, \dot{q}) \to \pi\dot{q} - H(\pi, q), \;\; \pi \equiv \frac{\partial L(q, \dot{q})}{\partial \dot{q}}. \quad (5.47)$$

Further, $\boldsymbol{H}(\pi, q)$ can be written in terms of creation and destruction operators, $\pi = -i(ae^{-i\phi} - a^\dagger e^{i\phi})\sqrt{\kappa/2}$, $q = (ae^{-i\phi} + a^\dagger e^{i\phi})/\sqrt{2\kappa}$, where κ and ϕ are arbitrary. One can see that if π and q were turned into operators they would obey the commutation rules,

$$[q_{\rm op}, \pi_{\rm op}] = \frac{1}{2}[(ae^{-i\phi} + a^\dagger e^{i\phi}), -i(ae^{-i\phi} - a^\dagger e^{i\phi})] = i. \tag{5.48}$$

There is a prescription for beginning with a Lagrangian, then finding the corresponding Hamiltonian written in terms of creation and destruction. Given that one would like these commutation relations hold true at all time, one may further constrain the choice of how to express q and π in terms of a and $a^\dagger$. The condition of satisfying the equations of motion will determine κ and ϕ (for the one-mode problem).
The process of beginning with a Lagrangian written in terms of fields, then writing the equivalent Hamiltonian expressing the field operators in terms of creation and destruction operator is referred to as "canonical quantization". The steps are:

- Beginning with $\boldsymbol{L}(\phi_i, \dot{\phi}_i)$, derive the conjugate momenta, $\pi_i = \partial \boldsymbol{L}/\partial\dot{\phi}_i$.
- Write ϕ_i and π_i as field operators so that

  $$[\phi_i(t), \pi_j(t)] = i\delta_{ij},$$

  where $\phi_i(t)$ are solutions to the equations of motion. Write your Hamiltonian as $\boldsymbol{H} = \pi_i\dot{\phi}_i - \boldsymbol{L}$ in terms of the field operators, and you're done.

If you're working with field operators in coordinate space, the discrete index i goes to the continuous label r and the operators must instead obey the equal-time commutation relation,

$$[\phi(r,t), \pi(r',t)] = i\delta^3(r - r'). \tag{5.49}$$

EXAMPLE:
Find the field operators and the Hamiltonian for the real field ϕ derived from the Lagrangian,

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

First find the conjugate momenta,

$$\pi(r) = \frac{\partial\mathcal{L}}{\partial\partial_t\phi(r)} = \partial_t\phi(r).$$

Now, note that if you write (from a guess)

$$\phi(r,t) = \int \frac{d^3p}{(2\pi)^3 2E_p}\left(e^{-ip\cdot r}a(p) + e^{ip\cdot r}a^\dagger(p)\right),$$

with $[a(p), a^\dagger(q)] = 2\pi(2\pi)^3\delta^3(p-q)$, this will also give

$$\pi(r,t) = -i\int \frac{d^3p}{(2\pi)^3 2E_p}\left(E_p e^{-ip\cdot r}a(p) - E_p e^{ip\cdot r}a^\dagger(p)\right).$$

Commuting ϕ with π at equal times will indeed give a delta function (multiplied by i) with a little bit of work. This demonstrates why the particular form for ϕ was chosen, and how if we had guessed a different expression for ϕ, it would have failed. For instance, if we had added an extra factor of F in the definition for ϕ, the commutation relation would have been off by a factor of F^2.

For the Klein-Gordon Lagrangian, the Hamiltonian is easy to derive

$$H = \int d^3r \left[\pi(r)\dot{\phi}(r) - \mathcal{L}\right] = \int d^3r \left[\frac{1}{2}\partial_t\phi(r,t)\partial_t\phi(r,t) + \frac{1}{2}\nabla\phi(r)\cdot\nabla\phi(r) + \frac{1}{2}m^2\phi(r)^2\right].$$

In a homework problem, you will show how this yields the expected expression when put in terms of creation and destruction operators.

Although it appears that path integrals allow one to equate quantum behavior to that of classical fields, there is an important difference. For classical fields, one uses the relation $\delta S = 0$, where the action S is $\int d^4x\mathcal{L}$. In the quantum system, one considers all paths, and adds their contribution coherently. Only in the limit $\hbar \to 0$, can one state that you can ignore all paths except the one path determined by minimum action. Thus, for quantum fields you can derive equations of motion for a field ϕ, but that does not mean that ϕ is determined. One can only determine the equation of motion for the evolution of the field averaged over the various paths, $\langle\phi\rangle$. This is similar to the harmonic oscillator, where you can work in the Heisenberg representation and solve for $\langle X(t)\rangle$. For example, in the ground state, one could imagine a system where $\langle X(t)\rangle = 0$ for all t, however if one were to measure the position, one would find a distribution of values, all non-zero but centered around zero.

5.7 Phonons and Fields

A lattice can be considered as being composed of coupled harmonic oscillators. A one-dimensional system could have masses m whose locations are x_n, which in a relaxed state are separated by a distance a. For harmonic oscillators, one can define a position operator in terms of the raising and lowering operators, a_n and $a_n^\dagger$. This operator, we will refer to as $\phi_n = x_n - na$, describes how far the lattice has moved from its relaxed state, and can be considered a field operator. The potential energy of the system due to two neighboring sites moving relative to one another is

$$V_n = \frac{1}{2}m\omega_0^2(\phi_{n+1} - \phi_n)^2, \tag{5.50}$$

where $m\omega^2$ is the spring constant. The kinetic energy of the n^{th} site is

$$K_n = \frac{1}{2}m\dot{\phi}_n^2. \tag{5.51}$$

Once can consider $\phi(x) = \phi_{n=x/a}$, then for small ϕ the total energy can be written as an integral over x,

$$E = \int dx \left\{\frac{a}{2}m\omega_0^2\left(\frac{\partial\phi}{\partial x}\right)^2 + \frac{m}{2a}\left(\frac{\partial\phi}{\partial t}\right)^2\right\}. \tag{5.52}$$

At this point, we can define a field

$$\chi \equiv \frac{\sqrt{m}}{a}\phi, \tag{5.53}$$

so that the energy becomes

$$E = \int dx \left\{ \frac{a^2\omega_0^2}{2}\left(\frac{\partial\chi}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial\chi}{\partial t}\right)^2 \right\}. \tag{5.54}$$

Aside from the factor of $a^2\omega_0^2$, this looks just like the Klein-Gordon equation for a real scalar field. This factor simply changes the speed of the propagation from c to $a\omega_0$, which is precisely what one expects for the speed of sound.
Thus, sound is composed of quantum excitations called phonons. Before the coupling there were N harmonic oscillator modes, and after the coupling, this switched to N modes in momentum space. Whereas the K.G. equation has modes for arbitrarily large momenta, or equivalently arbitrarily small wavelengths, phonons exist only until the N modes accounted for,

$$N = \frac{L}{2\pi}\int_{-k_D}^{k_D} dk = \frac{Lk_D}{\pi}. \tag{5.55}$$

Here, the cutoff in wave number is the Debye momentum, $k_D = \pi(N/L) = \pi/a$ and is proportional to the density. The more common expression for three dimensions is a bit more complicated. Equivalently, one can see that the Debye wavelength is related to the lattice spacing,

$$\lambda_D = \frac{2\pi}{k_D} = 2a. \tag{5.56}$$

5.8 Homework Problems

1. Consider the non-relativistic case presented in Sec. 5.1. Consider a state at $t = 0$ whose state is a wave packet of the form,

$$|k_0, \kappa\rangle = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{(2\pi\kappa^2)^{3/4}}\, e^{-(k-k_0)^2/4\kappa^2} a^\dagger(k)|0\rangle.$$

   (a) Calculate $\langle k_1, \kappa | k_2, \kappa\rangle$.

   (b) Calculate $\langle k_0, \kappa | H | k_0, \kappa\rangle$, where H is given in Eq. (5.12).

2. Consider a transformation to a frame moving with rapidity η along the x axis. For such a boost the Lorentz transformation matrix is

$$L^\alpha_{\ \beta} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

   Calculate the Jacobian for Lorentz transforming d^4x to d^4x'.

3. Express an eigenstate of the number operator, $|m\rangle$, in terms of coherent states, i.e., find $g_m(\eta)$ in the expression below,

$$|m\rangle = \int d\eta_r d\eta_i \, g_m(\eta)|\eta\rangle.$$

4. Consider the Hamiltonian derived at the end of 5.6,

$$H = \int d^3r \left[\pi(r)\dot{\phi}(r) - \mathcal{L}\right] = \int d^3r \left[\frac{1}{2}\partial_t\phi(r,t)\partial_t\phi(r,t) + \frac{1}{2}\nabla\phi(r)\cdot\nabla\phi(r) + \frac{1}{2}m^2\phi(r)^2\right].$$

   Using the expressions for the field operators in terms of $a(p)$ and $a^\dagger(p)$, derive an expression for H as an integral over momentum with the creation and destruction operators in the integrand.

5. The Lagrangian,

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

   with $\lambda = 1$ is known as the "Feynmann Gauge" for electromagnetic fields.

   (a) In the Feynmann gauge show that equivalently $\mathcal{L} = \frac{1}{2}A^\mu\partial^2 A_\mu$, i.e., show that the expressions integrated over d^4x are equal.

   (b) In terms of A and derivatives of A, what are the conjugates to the fields A^μ for the purposes of canonical quantization.

   (c) Using the expression for the fields in terms of creation and annihilation operators described in Eqs. (5.24) and (5.25), show that the conjugate fields indeed commute with the fields to create δ functions as required for canonical quantization. This constraint was responsible for the expressions in Eqs. (5.24) and (5.25).

6 Field Theory for Fermions

6.1 Beginning from the Lagrangian

One can follow the same procedure for canonical quantization as was used for bosons. Beginning with the Lagrangian,

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi, \tag{6.1}$$

the conjugate momenta are

$$\pi(r) = \frac{\partial \mathcal{L}}{\partial \partial_t \psi(r)} = i\bar{\psi}(r)\gamma^0 = i\psi^\dagger(r). \tag{6.2}$$

However, since the particles are fermions we look for a form for ψ that satisfies the anti-commutation laws,

$$\{\psi_i(r,t), i\psi_j^\dagger(r',t)\} = i\delta^3(r-r')\delta_{ij}. \tag{6.3}$$

A choice is

$$\psi(x) = \sum_s \int \frac{m d^3p}{E_p(2\pi)^3} \left(b_s(p) e^{-ip\cdot x} u_s(p) + d_s^\dagger(p) e^{ip\cdot x} v_s(p) \right), \tag{6.4}$$

where $b_s(p)$ and $d_s^\dagger(p)$ destroy particles of spin s and momentum p or create the corresponding anti-particles. They obey the anti-commutation laws,

$$\begin{aligned}
\{b_s(p), b_{s'}^\dagger(p')\} &= \frac{E_p}{m}(2\pi)^3\delta^3(p-p')\delta_{ss'}, \qquad (6.5)\\
\{d_s(p), d_{s'}^\dagger(p')\} &= \frac{E_p}{m}(2\pi)^3\delta^3(p-p')\delta_{ss'},\\
\{b_s(p), b_{s'}(p')\} &= \{d_s(p), d_{s'}(p')\} = \{b_s^\dagger(p), b_{s'}^\dagger(p')\} = \{d_s(p), d_{s'}(p')\} = 0\\
\{b_s(p), d_{s'}^\dagger(p')\} &= \{d_s(p), b_{s'}^\dagger(p')\} = 0.
\end{aligned}$$

Here the vectors $u_s(p)$ and $v_s(p)$ are normalized as:

$$\begin{aligned}
\bar{u}_s(p)u_{s'}(p) &= \delta_{ss'},\ \bar{v}_s(v)v_{s'}(p) = \delta_{ss'}, \qquad (6.6)\\
\bar{u}_s(p)v_{s'}(p) &= \bar{v}_s(p)u_{s'}(p) = 0.
\end{aligned}$$

To verify Eq. (6.4), one can anti-commute $\psi(x)$ and $\psi^\dagger(x')$ at equal times,

$$\begin{aligned}
\left\{\psi_i(x), \psi_j^\dagger(x')\right\} &= \sum_s \int \frac{m d^3p}{E_p(2\pi)^3} \left[e^{-ip\cdot(x-x')} u_{s,i}(p) u_{s,j}^\dagger(p) + e^{ip\cdot(x-x')} v_{s,i}(p) v_{s,j}^\dagger(p) \right] \Big|_{x_0=x_0'} \qquad (6.7)\\
&= \sum_s \int \frac{m d^3p}{E_p(2\pi)^3} e^{ip\cdot(x-x')} \left[u_{s,i}(p) u_{s,j}^\dagger(p) + v_{s,i}(-p) v_{s,j}^\dagger(-p) \right],
\end{aligned}$$

To proceed, note that $\sum_s u_{s,i}(p)\bar{u}_{s,j}(p)$ is a projection operator, that projects out the positive energy solutions. From a previous chapter, we know this to be (along with the one for negative energy solutions),

$$\sum_s u_{s,i}(p)\bar{u}_{s,j}(p) = \frac{\not{p}_{ij} + m\delta_{ij}}{2m}, \quad \sum_s v_{s,i}(p)\bar{v}_{s,j}(p) = \frac{\not{p}_{ij} - m\delta_{ij}}{2m}. \tag{6.8}$$

After one multiplies the two expressions above by γ^0, one can insert into the previous expression for the anti-commutator.

$$\sum_s u_{s,i}(p)u^\dagger_{s,j}(p) = \frac{E_p\delta_{ij} - \alpha_{ij}\cdot p + m\gamma^0_{ij}}{2m}, \quad \sum_s v_{s,i}(-p)v^\dagger_{s,j}(-p) = \frac{E_p\delta_{ij} + \alpha_{ij}\cdot p - m\gamma^0_{ij}}{2m}. \tag{6.9}$$

Each projector gives three contributions. The terms in the numerator proportional to m and $\alpha\cdot p$ disappear because the contributions from the positive- and negative-energy pieces cancel. The only term from the projector remaining is proportional to E_p/m. This factor cancels the identical factor in the invariant phase-space element. The remaining term is then seen to reproduce the delta functions in Eq. (6.3).

The current, which can also be derived from the Lagrangian is

$$\begin{aligned} j^\mu(x) &= \bar\psi(x)\gamma^\mu\psi(x) \qquad (6.10)\\ &= \sum_{ss'}\int \frac{md^3p}{E_p(2\pi)^3}\frac{md^3p'}{E_{p'}(2\pi)^3} \\ &\quad \Big[b^\dagger_s(p)b_{s'}(p')\bar u_s(p)\gamma^\mu u_{s'}(p')e^{i(p-p')\cdot x} + d_s(p)d^\dagger_{s'}(p')\bar v_s(p)\gamma^\mu v_{s'}(p')e^{-i(p-p')\cdot x} \\ &\quad + b^\dagger_s(p)d^\dagger_{s'}(p')\bar u_s(p)\gamma^\mu v_{s'}(p')e^{i(p+p')\cdot x} + d_s(p)b_{s'}(p')\bar v_s(p)\gamma^\mu u_{s'}(p')e^{-i(p+p')\cdot x}\Big] \end{aligned}$$

The cross terms above are important, as they are related to pair creation when the current is coupled to an electromagnetic field, A_μ. However, if one integrates over all x, one finds delta functions in the momentum. The net charge is then

$$\begin{aligned} Q = \int d^3x j^{\mu=0}(t=0,x) \;&= \qquad (6.11)\\ \sum_{ss'}\int \frac{m^2d^3p}{E_p^2(2\pi)^3} &\left[b^\dagger_s(p)b_{s'}(p)u^\dagger_s(p)u_{s'}(p) + d_s(p)d^\dagger_{s'}(p)v^\dagger_s(p)v_{s'}(p)\right. \\ &\left. + b^\dagger_s(p)d^\dagger_{s'}(-p)u^\dagger_s(p)v_{s'}(-p) + d_s(p)b_{s'}(-p)v^\dagger_s(p)u_{s'}(-p)\right]. \end{aligned}$$

The overlaps above can be calculated (see HW problems),

$$\begin{aligned} u^\dagger_s(p)u_{s'}(p) &= v^\dagger_s(p)v_{s'}(p) = \frac{E_p}{m}\delta_{ss'}, \qquad (6.12)\\ u^\dagger_s(p)v_{s'}(-p) &= v^\dagger_s(-p)u_{s'}(p) = 0. \end{aligned}$$

This results in

$$Q = \sum_s\int\frac{md^3p}{E_p(2\pi)^3}\left[b^\dagger_s(p)b_s(p) + d_s(p)d^\dagger_s(p)\right]. \tag{6.13}$$

Anti-commuting the operators d and $d^\dagger$ gives

$$Q = \sum_s\int\frac{md^3p}{E_p(2\pi)^3}\left[b^\dagger_s(p)b_s(p) - d^\dagger_s(p)d_s(p) + \text{infinite constant}\right] \tag{6.14}$$

The same infinite constant was found for the boson case considered in the previous chapter, only this background charge has the opposite sign. Aside from this constant, the result is pretty much what was expected: the number of positive minus the number of negative particles.

6.2 Pair Creation

One of the most common problems one faces in physics is pair creation, often the production of a particle-antiparticle pair. Any classical field that couples to $\bar{\Psi}\Psi$, or $\bar{\Psi}\gamma^\mu\Psi$, or similar coupling, will involve particle creation. In this section, we consider an example using the "sudden" approximation in quantum mechanics, where the vector potential suddenly changes values. Similarly, one could consider the mass suddenly changing. One motivation for working through this example is that it underscores the meaning and implication of negative-energy solutions to the Dirac equation.

EXAMPLE:
We will calculate pair creation from a pulse of electric field. Consider an electric field that is constant in space but exists only instantaneously,

$$\vec{E}(r,t) = \alpha\hat{z}\delta(t).$$

Find the density of particle and anti-particle pairs of momentum $k = k_z\hat{z}$ and $-k_z\hat{z}$, created by the pulse.

For the vector potential, choose

$$A_z(t) = \alpha\Theta(t).$$

For $t < 0$ we will look at the solution to the Dirac equation $v_s(-k)e^{iE_kt+ik\cdot r}$. This represents the negative energy solution, which is filled. Evolving the state beyond $t = 0$, one can find the overlap with the state $u'_s(k)e^{-iE'_kt+ik\cdot r}$. This represents the disappearance of the filled negative-energy state through the creation of a particle of momentum k. The destruction of a negative-energy state of momentum k can be thought of as the creation of a positive-energy anti-particle with momentum $-k$. Thus the squared overlap is the probability that the mode goes from zero particles to a pair.

To proceed, we must first express the wave functions $v_s(k)$ and $u'_{s'}(k)$. The prime refers to the fact that u will be evaluated for $t > 0$, which has a different vector potential. Recall that adding a vector potential means that $u(k, A \neq 0)$ will be equal to $u(k-eA, A = 0)$. Since k are good quantum numbers, and since we are interested in all spins, we can write the transition probability as

$$\mathcal{P} = \sum_s \left| u_s^\dagger(k_z, A_z)v_s(-k_z, \vec{A} = 0)\right|^2. \tag{6.15}$$

We only consider the same-helicity elements because the matrix element would be zero otherwise. If we were to perform the calculations for k not along the z axis, we would have to extend the calculation to more spin combinations.

First, we write the spinors as:

$$\begin{aligned} u_s(k_z) &= \frac{1}{\sqrt{2E'_k(E'_k+m)}}(E'_k\gamma^0 - k'_z\gamma_z + m)u_s(0), \\ v_s(-k_z) &= \frac{1}{\sqrt{2E_k(E_k+m)}}(-E_k\gamma^0 - k_z\gamma_z + m)v_s(0). \end{aligned} \tag{6.16}$$

These expresssions differ from what was derived in a previous HW assignment by a factor $\sqrt{m/E_k}$. With this choice, $u^\dagger(k)u(k) = 1$, rather than E_k/m. One can calculate the overlap by inspection by looking at the Dirac matrices and realizing that one need only consider terms that mix upper and lower components (defined by Dirac representation).

$$\mathcal{P} = \frac{1}{[2E_k(E_k+m)][2E'_k(E'_k+m)]}\left\{(E'_k+m)k_z - k'_z(E_k+m)\right\}^2 \qquad (6.17)$$

Here $k'_z = k_z - e\alpha$ and $E'_k = \sqrt{k'^2_z + m^2}$. The above probability refers to a single mode, and should be multiplied by 2 to account for spins. To get $dN_{\rm pairs}/d^3k$, one must add a factor of $V/(2\pi)^3$ for phase space.

EXAMPLE:
Consider the same problem as above, only solve with a Bogoliubov transformation.

Consider the field operators before and after the change in $\vec{A}$,

$$\begin{aligned}
\Psi(r, t=0^-) &= \sum_s \int \frac{m d^3k}{E_k(2\pi)^3} e^{ik\cdot r}\left\{b_s(k)u_s(k) + d^\dagger_s(-k)v_s(-k)\right\}, \qquad (6.18)\\
\Psi'(r, t=0^+) &= \sum_s \int \frac{m d^3k}{E'_k(2\pi)^3} e^{ik\cdot r}\left\{b'_s(k)u'_s(k) + d'^\dagger_s(-k)v'_s(-k)\right\}.
\end{aligned}$$

The two operators must be equal at time $t = 0$. Further, one can project out $b_s(k)$ from the first relation using the orthogonality relations involving $u_s(k)$ and $v_s(-k)$ from the homework.

$$b'_{s'}(k) = \int d^3r\, e^{-ik\cdot r} u'^\dagger_{s'}(k)\Psi'(r, t=0). \qquad (6.19)$$

Since the two field operators are the same at $t = 0$, we can also write $b'_{s'}(k)$ in terms of the overlap with $\Psi(r, t=0)$,

$$\begin{aligned}
b'_{s'}(k) &= \int d^3r\, e^{-ik\cdot r} u'^\dagger_{s'}(k)\Psi(r, t=0) \qquad (6.20)\\
&= \sum_s \int \frac{m d^3k'}{E_k(2\pi)^3}\int d^3r\, e^{-i(k-k')\cdot r}\left\{u'^\dagger_{s'}(k)u_s(k')b_s(k') + u'^\dagger_{s'}(k)v_s(-k')d^\dagger_s(-k')\right\}\\
&= \sum_s \frac{m}{E_k}\left\{u'^\dagger_{s'}(k)u_s(k)b_s(k) + u'^\dagger_{s'}(k)v_s(-k)d^\dagger_s(-k)\right\}
\end{aligned}$$

This re-expression of the creation and destruction operators is known as Bogoliubov transformation. The state at $t = 0$ is the vacuum in the b, $b^\dagger$ basis, i.e., $b(k)|0\rangle = 0$, $b'(k)|0\rangle \neq 0$. If one calculates $\langle 0|b'^\dagger_{s'}(k)b'_{s'}(k)|0\rangle$, one will get contributions from the

$d^\dagger$ term above. This gives

$$\begin{aligned}(2\pi)^3\frac{E_k dN}{m d^3k} &= \langle 0|b'^\dagger_{s'}(k)b'_{s'}(k)|0\rangle \qquad (6.21)\\ &= \sum_s \frac{m^2}{E_k^2}|u'^\dagger_{s'}(k)v_s(-k)|^2\langle 0|d_s(k)d^\dagger(k)|0\rangle \\ &= (2\pi)^3\frac{m}{E_k}\delta^3(k-k)\sum_s |u'^\dagger_{s'}(k)v_s(-k)|^2 \\ &= V\frac{m}{E_k}\sum_s |u'^\dagger_{s'}(k)v_s(-k)|^2, \\ \frac{dN}{d^3k} &= \frac{V}{(2\pi)^3}\left(\frac{m}{E_k}\right)^2\sum_s |u'^\dagger_{s'}(k)v_s(-k)|^2. \qquad (6.22)\end{aligned}$$

Since this calculation was done with the normalization $u^\dagger u = E/m$, this is exactly the same result as above. Note that summing over modes brings out the factor $V/(2\pi)^3$.

Bogoliubov transformations turn out to be useful for a variety of problems, most famously superconductivity. They tend to be useful for problems where an interaction or perturbation is of the form,

$$V = \sum_k V_k(b_k^\dagger d_{-k}^\dagger + b_k d_{-k}). \qquad (6.23)$$

For superconductivity, these would be interactions that create or destroy Cooper pairs. The same form can occur in many instances in relativistic treatments. For instance, if an interaction adds to a particle's mass,

$$V = \int d^3r\, \delta M(r)\bar\Psi(r)\Psi(r), \qquad (6.24)$$

one can expand the field operators, which will lead to terms of the type mentioned above.

6.3 Homework Problems

1. Using the fact that $u_s(p) = (\not{p} + m)u_s(p=0)/\sqrt{2m(E_p+m)}$ and $v_s(p) = (-\not{p} + m)v_s(p=0)/\sqrt{2m(E_p+m)}$ (where the solutions $u_s(p)$ and $v_s(p)$ are orthonormalized), show that:
   - $\bar u_s(p)u_{s'}(p) = \delta_{ss'}$
   - $\bar u_s(p)v_{s'}(p) = 0$
   - $u_s^\dagger(p)u_{s'}(p) = \delta_{ss'}(E_p/m)$
   - $u_s^\dagger(p)v_{s'}(-p) = 0$

   HINT: Remember that $\vec\gamma$ are anti-Hermitian.

2. Derive T_{00} from the Lagrangian for free fermions, $\mathcal{L} = \bar{\psi} i \not{\partial} \psi - m \bar{\psi} \psi$,

$$T_{00} = \frac{\partial \mathcal{L}}{\partial \dot{\Psi}} \dot{\Psi} - \mathcal{L}.$$

   Then express $H = \int d^3 r \, T_{00}$ in terms of integrals over creation and destruction operators in momentum space.

3. Consider the quantity $\bar{u}_s(p) v_{s'}(-p)$. Terms like this arise in the case where a scalar field ϕ might couple to particle-antiparticle production. Assuming p lies along the z axis, calculate the overlap. Express your answer in terms of E_p, m and p_z.

4. Consider the quantity $\bar{u}_s(p) \gamma^\mu v_{s'}(-p)$. Terms like this arise in the case where a vector field A^μ might couple to particle-antiparticle production. Assuming p lies along the z axis, calculate the overlap for all four values of μ, and for all combinations of s and s', where the spin labels refer to various eigenstates of Σ_z. Express your answer in terms of E_p, m and p_z.

7 Propagators

7.1 Why Propagators

Time dependent perturbation theory has played the dominant role in the phenomenology of particle physics over the last several decades. The changes to the evolution operator in such a theory look like

$$\cdots + \frac{(-i)^n}{n!}\int dt_1 \cdots dt_n \mathcal{T}\left[V(t_1)\cdots V(t_2)\right] + \cdots \,, \tag{7.1}$$

where $\mathcal{T}$ is the time ordering operator. The elements $V(t_1)$ are the perturbations in the interaction representation. The perturbations V involve integrals over creation and destruction operators, which in the interaction representation become $a_k^\dagger(t) = e^{iH_0 t}a_k^\dagger e^{-iH_0 t} = e^{iE_k t}a_k^\dagger$. Matrix elements are thus constructed from terms of the form $\langle 0|a_k(t_1)a_{k'}^\dagger(t_2)|0\rangle$. Of course, the interactions are usually expressed in terms of field operators in coordinate space, which means that the quantities of interest are $\mathcal{T}\langle 0|\Phi(x',t')\Phi(x,t)|0\rangle$. These are referred to as time-ordered propagators, and perturbation theory is generally built from these quantities. They are also known as Green's functions or Feynmann propagators.
We will begin by finding an expression for the Feynmann propagator for real scalar fields. Once these are known, propagators for other particles are often expressed as a product of the scalar propagator and additional terms that carry the spin information. The scalar propagator $G_F(x)$ is

$$\begin{aligned} G_F(x) &\equiv i\mathcal{T}\langle 0|\Phi(x)\Phi(0)|0\rangle \\ &= i\int \frac{d^3k d^3q}{(2\pi)^6(2E_k 2E_q)}\,\mathcal{T}\langle 0|\left[a(k)e^{-ik\cdot x} + a^\dagger(k)e^{ik\cdot x}\right]\left[a(q)+a^\dagger(q)\right]|0\rangle \\ &= i\Theta(x_0)\int \frac{d^3k}{(2\pi)^3 2E_k}e^{-ik\cdot x}\Big|_{k_0=E_k} + \Theta(-x_0)\int \frac{d^3k}{(2\pi)^3 2E_k}e^{ik\cdot x}\Big|_{k_0=E_k}. \end{aligned} \tag{7.2}$$

From contour integration, one can see that

$$G_F(x) = \int \frac{d^4k}{(2\pi)^4}e^{-ik\cdot x}\frac{-1}{k^2 - m^2 + i\epsilon},\ \epsilon \to 0 \tag{7.3}$$

The term $-1/(k^2-m^2+i\epsilon)$ is the Fourier transform of $G_F(x)$, and is known as $G_F(k)$. The motivation for using the propagator in momentum space derives from perturbation theory, where most expressions involve integrations of products of propagators over all time, and integrated over all space. These integrals can usually be expressed more naturally in momentum and frequency space. This is certainly the case for scattering problems, which are based on adiabatic pictures involving wave packets slowly passing over a large time.

7.2 Propagator for Massless Scalar Field

There are instances, such as deriving Coulomb's law, which do involve knowing the propagator in coordinate space. For that reason, we perform the integration listed above to explicitly solve

for $G_F(x)$ for the example of massless particles. The derivation is surprisingly painful:

$$\begin{aligned} G_F(x) &= i\int \frac{d^3k}{(2\pi)^3 2k}\left[e^{-ik\cdot x}\Theta(x_0)+e^{ik\cdot x}\Theta(-x_0)\right]\big|_{k_0=E_k} \\ &= \frac{i}{8\pi^2}\int_{-1}^{1} d\cos\theta \int_0^\infty k dk\ \left[e^{-ikx_0+ikr\cos\theta}\Theta(x_0)+e^{ikx_0-ikr\cos\theta}\Theta(-x_0)\right] \\ &= \frac{i}{4\pi^2 r}\int_0^\infty dk\left[e^{-ikx_0}\sin(kr)\Theta(x_0)+e^{ikx_0}\sin(kr)\Theta(-x_0)\right] \end{aligned} \tag{7.4}$$

Here, r is the magnitude of the spatial components of x. To proceed, we will consider the real and imaginary parts separately, $G_F = G_{F,r} + iG_{Fi}$. For the imaginary part,

$$\begin{aligned} G_{F,i}(x) &= \frac{1}{4\pi^2 r}\int_0^\infty dk\ \cos(kx_0)\sin(kr) \\ &= \frac{1}{16\pi^2 ri}\int_0^\infty dk\ \left[e^{ik(x_0+r)}-e^{-ik(x_0+r)}+e^{-ik(x_0-r)}-e^{ik(x_0-r)}\right] \\ &= \frac{-1}{8\pi^2 r}\left[\frac{1}{x_0+r}-\frac{1}{x_0-r}\right] \\ &= \frac{-1}{4\pi^2 x^2}. \end{aligned} \tag{7.5}$$

To calculate the real part,

$$\begin{aligned} G_{F,r} &= \frac{1}{4\pi^2 r}\int_0^\infty dk\ \sin(k|x_0|)\sin(kr) \\ &= \frac{-1}{16\pi^2 r}\int_0^\infty dk\ \left[e^{ik(|x_0|-r)}+e^{-ik(|x_0|-r)}\right] \\ &= \frac{-1}{16\pi^2 r}\int_{-\infty}^\infty dk\, e^{ik(|x_0|-r)} \\ &= \frac{-1}{8\pi r}\delta(|x_0|-r). \end{aligned} \tag{7.6}$$

Adding the two pieces together,

$$G_F(x) = \frac{-i}{4\pi^2}\cdot\frac{1}{x^2-i\epsilon}. \tag{7.7}$$

7.3 Propagator for the electromagnetic field

The propagator for the electromagnetic field is easily calculated in the Feynmann gauge. In that gauge, the field operators appear as four different operators, but with the zero$^{\text{th}}$ component commuting with the wrong sign. Thus, the propagator is

$$\begin{aligned} G^{\mu\nu}(x) &= \mathcal{T}\langle 0|A^\mu(x)A^\nu(0)|0\rangle = -ig^{\mu\nu}G_F(x) \\ &= -g^{\mu\nu}\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2+i\epsilon} = \frac{g^{\mu\nu}}{4\pi^2}\cdot\frac{1}{x^2-i\epsilon} \end{aligned} \tag{7.8}$$

Finally, with all the mathematical apparatus compiled so far, we are able to calculate Coulomb's law. From perturbation theory, the impact of a field A^μ coupled to the static charge density, $V = \int d^3r A^\mu j_\mu$ with only zero$^{\text{th}}$ components, is

$$\begin{aligned}
|\psi\rangle &= |0\rangle - i\int_{-\infty}^{0} dt\, V(t)|0\rangle, \qquad (7.9)\\
\langle\psi| &= \langle 0| + i\langle 0|\int_{-\infty}^{0} dt\, V(t),\\
\langle\psi|V(0)|\psi\rangle &= 2\int d^3r d^3r' \int_{-\infty}^{0} dt\, j_0(r) j_0(r')\, \mathrm{Imag}\left\{\langle 0|A^0(r',0)A^0(r,t)|0\rangle\right\}\\
&= 2\int d^3r d^3r' \int_{0}^{\infty} dt\, j_0(r) j_0(r')\, \mathrm{Imag}\left\{G^{00}(r-r',t)\right\},\\
&= \frac{1}{8\pi}\int d^3r d^3r'\, j_0(r) j_0(r') \frac{1}{|r-r'|},
\end{aligned}$$

where the earlier expression for the imaginary propagator was inserted. The expression is exactly what was expected all along, but the derivation seems surprisingly involved. Additionally, this emphasizes the role of the "unphysical" polarization modes, i.e., those with the wrong-signed commutation relations.

Similarly, one can calculate the potential energy for a scalar density $s(r)$ interacting with another scalar density, $s(\mathrm{r}')$, due to a coupling with a scalar field, $V = \int d^3r\, s(r)\Phi(\mathrm{r})$. The propagator is identical to the electromagnetic case above, apart from the term $-g^{\alpha\beta}$. One finds

$$\langle\psi|V(0)|\psi\rangle = \frac{-1}{8\pi|r-r'|}\int d^3r d^3r'\, s(r)s(r'). \qquad (7.10)$$

Whereas the electromagnetic interaction between charges of the same sign are repulsive, the interactions between scalar "charges" of the same sign are attractive. This plays an important role in nuclear physics. Nuclei are made of protons and neutrons. The neutrons couple to one another four different ways:

- Exchange of the ω meson. The ω is a massive vector meson that couples to baryon number. Hence, both protons and neutrons have the same sign. Since it is a vector interaction, the force is repulsive.
- Exchange of the σ meson. This is a proxy for two-pion exchange, and is a scalar potential that couples to massive protons and neutrons identically. However, since this is a scalar field, the interaction is attractive.
- Exchange of the ρ_0 meson. The ρ couples to neutrons and protons with opposite sign. Since there are roughly equal numbers of neutrons and protons, this interaction tends to be smaller than that from the ω meson. This interaction largely cancels when the number of neutrons and protons are equal. Whenever there is a large isospin asymmetry (more of one species of nucleon than the other) this interaction becomes repulsive, and is responsible for a significant fraction of the nuclear symmetry energy.
- Single pion exchange. Pions are pseudo scalars, and their couplings change according to the spin of the particle. Additionally, they carry isospin like the ρ_0. Thus, they are doubly inclined to have their interactions cancel. Even if the neutron and proton fractions were to

be notably different, the interaction would cancel if the number of spin-up and spin-down protons (or neutrons) were equal. Thus, even though the pion is the lightest meson, thus making such an interaction less screened than the others, this cancellation makes it less important than the first two. The spin-spin nature of the interaction is often referred to as the "tensor" interaction. As one might guess, it plays an important role in driving the forces responsible for pairing.

The attractive force due to σ exchange, and the repulsive force due to ω exchange would each contribute several hundreds of MeV to the nuclear potential if considered singularly. However, after the cancellation, the net depth of the nuclear potential is only ≈ 35 Whereas the forces cancel, the corresponding relativistic corrections, a.k.a. the spin-orbit interaction, come in with the same sign. Thus, the spin-orbit interaction in nuclear physics is approximately a factor of twenty greater in magnitude than expectations assuming only a scalar or only a vector interaction.

7.4 Causality

Surprisingly, proving that a theory is causal is more difficult that one might first guess. For a massless scalar theory, one can consider the propagation of a field that is altered by a source function, $s(x), x_0 < 0$, and then recorded at $x = 0$. To first order in perturbation theory,

$$\delta\langle\Phi(x=0)\rangle = \frac{-i}{\hbar}\int_{-\infty}^{0} dt\langle[\Phi(0), V(t)]\rangle, \tag{7.11}$$

$$V = \int d^3x\, j(x)\Phi(x). \tag{7.12}$$

This can be rewritten using in terms of the propagator defined in Eq. (7.7),

$$\delta\langle\Phi(x=0)\rangle = \mathcal{R}e\int_{x_0<0} d^4x\, G_F(x)j(x) \tag{7.13}$$

$$= \int_{x_0<0} d^4x\, j(x)\mathcal{R}e\frac{-i}{4\pi^2}\frac{1}{x^2 - i\epsilon} \tag{7.14}$$

$$= \frac{1}{4\pi}\int_{x_0<0} d^4x\, j(x)\delta(x^2). \tag{7.15}$$

Here, it is clear that massless particles propagate at the speed of light. With some more work, it can be shown that by adding more complicated intermediate steps to the evolution, causality is maintained as long as each time-ordered propagator is causal.

7.5 Massive Scalar Field

The real and imaginary parts of the propagator for massive scalar fields can be found as was done in Eq.s (7.5) and (7.6). The difference being that $E_k \neq k$, but instead equals $E_k =$

$\sqrt{k^2+m^2}$.

$$\begin{aligned} G_{F,r}(x) &= \frac{1}{(2\pi)^2}\int_0^\infty \frac{k^2 dk}{2E_k} d\cos\theta \left[e^{iE_k x_0 + ikr\cos\theta}\Theta(x_0) + e^{-iE_k x_0 + ikr\cos\theta}\Theta(-x_0)\right] \\ &= \frac{1}{8\pi^2 r}\int_0^\infty \frac{k dk}{E_k}\left[\sin\left(E_k x_0 + kr\right) - \sin\left(E_k x_0 - kr\right)\right] \\ G_{F,i}(x) &= \frac{1}{16\pi^2 r}\int_{-\infty}^\infty \frac{k dk}{E_k}\left[e^{iE_k x_0 - ikr} - e^{iE_k x_0 + ikr}\right] \end{aligned} \tag{7.16}$$

These don't collapse to a particularly transparent form, but when calculating the perturbative correction to the energy one does find a simple expression. As was done for Eq. (7.9),

$$\begin{aligned} \langle\psi|V(0)|\psi\rangle &= -2\int d^3r d^3r' \int_0^\infty dt\, s(r)s(r')e^{-\epsilon|t|}G_{F,r}(r-r',t) \\ &= \frac{-1}{4\pi^2}\int d^3r d^3r' \frac{1}{|r-r'|}s(r)s(r')\int \frac{k dk}{E_k^2 + i\epsilon}\sin(k|r-r'|) \\ &= \frac{-1}{8\pi}\int d^3r d^3r' \frac{e^{-m|r-r'|}}{|r-r'|}s(r)s(r') \end{aligned} \tag{7.17}$$

The potential, e^{-mr}/r, is known as the Yukawa potential. The exponential factor can be crudely motivated from the uncertainty principle, by stating that the intermediate state is off-shell by an amount m, thus allowing it to propagate a distance or time of $1/m$.

7.6 Propagators vs. Green's Functions

Propagators are of the form $i\mathcal{T}\langle\Phi(x)\Phi(0)\rangle$. Green's functions refer to a method for solving differential equations. The equations are typically of the form,

$$K\phi(t) = j(t), \tag{7.18}$$

where K is some function of derivatives. For instance, in the Klein-Gordon equation for a specific momentum mode, $K = -\partial_t^2 - E_k^2$. For a harmonic oscillator, where $\phi \to x$, $K = \partial_t^2 + (k/m)$. The term $j(t)$ is a driving term. For the harmonic oscillator this corresponds to an external force, whereas for the Klein-Gordon equation this corresponds to a source term of particles of a scalar form. For the electromagnetic case, it will appear as an electromagnetic current, which by coupling to the field will radiate particles. If one writes the differential equation in Fourier space,

$$K(\omega)\phi(\omega) = j(\omega), \tag{7.19}$$

where any derivatives inscribed in K become factors of ω, i.e., $i\partial_t \to \omega$. Thus, for the K.G. equation, $K(\omega) = -\omega^2 + E_k^2$.
Green's functions are solutions of the form,

$$\phi(t) = \int_{-\infty}^\infty dt'\, G_>(t-t')j(t'). \tag{7.20}$$

The subscript ">" denotes that this is a "retarded" Green's function and that $G(t>0) = 0$. This corresponds to the BC that ϕ would be zero for all times before the field is turned on. The

"advanced" Green's function is zero for $t < 0$ and corresponds to the BC where the $\phi = 0$ for times after the source is turned off. Note that to this point the fields are purely classical objects. In Fourier space,

$$\phi(\omega) = G(\omega) j(\omega). \tag{7.21}$$

Combined with the equations of motion in Fourier space, one can see that

$$G_>(\omega) = K^{-1}(\omega). \tag{7.22}$$

For the Klein-Gordon equation, it would seem that

$$G_>(\omega) = \frac{-1}{\omega^2 - E_k^2}? \tag{7.23}$$

However, there are poles for $\omega = \pm E_k$. In fact, one could have guessed at the same answer for $G_<$. To choose how to handle the poles, one needs to look at the BC. If $G(t) \sim \int d\omega e^{-i\omega t} G(\omega)$ is to be zero for negative times, $G_>(\omega)$ must have no poles in the upper half plane. The opposite must be true for $G <$. Thus,

$$G_>(\omega) = \frac{-1}{(\omega - i\epsilon)^2 - E_k^2}, \; G_<(\omega) = \frac{-1}{(\omega + i\epsilon)^2 - E_k^2}. \tag{7.24}$$

In comparison the Feynmann propagator is

$$G_F(\omega) = \frac{-1}{\omega^2 - E_k^2 + i\epsilon} = \frac{-1}{(\omega - E_k + i\epsilon)(\omega + E_k - i\epsilon)}. \tag{7.25}$$

Thus, the poles for the Feynmann propagator are split between the upper and lower half planes, with the positive frequency pole being below the plane.
Thus far, the similarity seems only in the form for the result. To see a deeper connection, one must consider Linear response theory. The derivation that follows forms the foundation of quantum response for condensed matter systems, and thus is well worth your time to follow.
In linear response theory, one considers to quantum operators $\phi(t)$ and $B(t)$, where one has an interaction term,

$$V_I(t) = B(t) F(t), \tag{7.26}$$

where F is an external field. For instance, A could be the current, A could be the charge density multiplied by x and F would be the electric field. Linear response theory solves for the coefficient, $\chi(t - t')$, such that

$$\delta\langle\phi(t)\rangle = \int_{-\infty}^{t} \chi(t - t') F(t'), \tag{7.27}$$

where the change of ϕ arises from the coupling of A to F. From perturbation theory,

$$\begin{aligned} \delta\langle\phi\rangle &= \frac{-i}{\hbar}\int_{-\infty}^{0} dt` \, \langle\phi(0) V_I(t')\rangle + \frac{i}{\hbar}\int_{-\infty}^{0} dt' \, \langle V_I(t')\phi(0)\rangle \\ &= \frac{-i}{\hbar}\int_{-\infty}^{0} dt' \, \langle[\phi(0), A(t')]\rangle F(t'). \end{aligned} \tag{7.28}$$

Now, one can consider the case where $V = j(t)\Phi(t)$, and the field $\phi = \Phi$. In the that case the equations of motion for ϕ have a source term j as seen above. Then, by inspection one can see that,

$$\begin{aligned} \langle \Phi(t) \rangle &= \int_{-\infty}^{t} dt' \, G_>(t-t') j(t'), \\ G_>(t-t') &= \frac{-i}{\hbar} \langle [\Phi(0), \Phi(t)] \rangle. \end{aligned} \tag{7.29}$$

Now, the connection between Green's functions, which seem to apply to classical theories, and quantum theories, which involves non-commuting operators, seems more clear. However, the fact that the $i\epsilon$s appear differently for the different propagators, makes it difficult to express $G_F(\omega)$ in terms of $G_>$ and $G_<$. In practice, you simply identify all the poles on the real axis, then force those poles above and below the real axis depending on whether they are positive or negative poles.

The source of all the ugliness associated with seemingly adding $i\epsilon$s to get the behavior one wants derives from the fact that the propagations are undamped. Physically, this means one must worry about boundary conditions for the fields at $t = \pm\infty$. If one adds damping, the poles move off the axis and everything becomes clear. In fact, $G_<$ goes to zero as there are no fields from $t = -\infty$ to annihilate.

7.7 Gauge Invariance

The Lagrangian for the electromagnetic field is

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{7.30}$$

For reasons of convenience, which are not at all readily apparent, we will add a "gauge-fixing" term,

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{\lambda}{2} (\partial \cdot A)^2 = \frac{-1}{2} \left[\partial_\alpha A^\mu \partial^\alpha A_\mu \right] - \frac{(\lambda - 1)}{2} (\partial \cdot A)^2. \tag{7.31}$$

For $\lambda = 1$, this looks like four independent fields, $A^0 \cdots A^3$, and leads to the propagator we used initially. This propagator featured four independent fields, though with the zero$^{\text{th}}$ component being a negative norm state. The choice of $\lambda = 1$ is known as the Feynmann gauge.

What we will show below is that results are independent of λ due to the fact that A couples only through a conserved current. We begin by dividing the propagator for $\lambda = 1$ into two pieces, and refer to it as $G^{\alpha\beta}_{\lambda=1}$

$$G^{\alpha\beta}_{\lambda=1}(k) = -i \frac{g^{\alpha\beta} - k^\alpha k^\beta / k^2}{k^2 + i\epsilon} - i \frac{k^\alpha k^\beta / k^2}{k^2 + i\epsilon}. \tag{7.32}$$

This can be written as two pieces, with each characterized by a projector,

$$G^{\alpha\beta}_{\lambda=1} = -i \frac{P_T(k)^{\alpha\beta} + P_L(k)^{\alpha\beta}}{k^2 + i\epsilon}, \tag{7.33}$$

$$P_T^{\gamma\beta}(k) = g^{\gamma\beta} - k^\alpha k^\beta / k^2, \;\; P_L^{\gamma\beta}(k) = k^\alpha k^\beta / k^2. \tag{7.34}$$

It is easy to see that these are projectors, and that $k_\alpha P_T(k)^{\alpha\beta} = 0$.
Switching gears, we can now consider the term $(\lambda-1)(\partial\cdot A)^2/2$ perturbatively. The propagator can be written as:

$$\begin{aligned} G^{\alpha\beta}(x-x') &= G^{\alpha\beta}_{\lambda=1}(x) \\ &\quad + \sum_{n=1}^{\infty} \frac{(-i)}{n!} \int d^4y_1 \cdots d^4y_n \, \mathcal{T}\langle 0|A^\alpha(x)V(y_1)\cdots V(y_n)A^\beta(x')|0\rangle, \\ V(y) &= \frac{(\lambda-1)}{2}(\partial_\gamma A^\gamma(y))(\partial_\delta A^\delta(y)). \end{aligned} \tag{7.35}$$

The next step requires some head scratching. One can see that the matrix elements all require each field to couple to some other field. Since there are $n!$ ways to connect the fields, and since each connection is identical, one can see that

$$\begin{aligned} G^{\alpha\beta}(x-x') &= G^{\alpha\beta}_{\lambda=1}(x-x') \\ &\quad + \sum_{n=1}^{\infty} [i(\lambda-1)]^n \int d^4y_1 \cdots d^4y_n \; [\partial_{\gamma 1}\partial_{\gamma 2} G^{\alpha\gamma_1}_{\lambda=1}(x-y_1)] \\ &\quad \cdots \left[\partial_{\gamma n-1}\partial_{\gamma n} G^{\gamma_{n-1}\gamma_n}_{\lambda=1}(y_{n-1}-y_n)\right] G^{\gamma_n\beta}_{\lambda=1}(y_n-x'). \end{aligned} \tag{7.36}$$

This is moderately less hideous in momentum space. After Fourier transforming,

$$\begin{aligned} G(k) &= G_{\lambda=1}(k) \\ &\quad + \sum_{n=1}^{\infty} [(\lambda-1)k^2]^n G_{\lambda=1}(k)P_L(k)G_{\lambda=1}(k)P_L(k)\cdots P_L(k)G_{\lambda=1}(k), \end{aligned} \tag{7.37}$$

where the fact that the two derivatives acting on each point could be identified as the projector, i.e. $k^\alpha k^\beta = k^2 P_L^{\alpha\beta}(k)$. Now, one can use the fact that

$$-k^2 P_L(k) G_{\lambda=1}(k) = P_L, \tag{7.38}$$

to obtain

$$\begin{aligned} G(k) &= P_T G_{\lambda=1}(k) + P_L G_{\lambda=1}(k) + P_L G_{\lambda=1}(k) \sum_{n=1}^{\infty} (1-\lambda)^n \\ &= P_T G_{\lambda=1}(k) + \frac{1}{\lambda} P_L G_{\lambda=1}(k). \end{aligned} \tag{7.39}$$

Including the Lorentz indices,

$$\begin{aligned} G^{\alpha\beta}(k) &= \frac{P_T^{\alpha\beta}}{k^2+i\epsilon} + \frac{P_L^{\alpha\beta}}{\lambda k^2 + i\epsilon} \\ &= \frac{g^{\alpha\beta}}{k^2+i\epsilon} + \frac{(1-\lambda)P_L^{\alpha\beta}}{\lambda k^2 + i\epsilon} \end{aligned} \tag{7.40}$$

For $\lambda = 1$, one returns to the Feynmann gauge. For $\lambda = \infty$, the longitudinal polarization contribution vanishes. This corresponds to the condition $\partial\cdot A = 0$, and is known as the Lorentz

gauge. It should not be surprising that by adding an infinite term proportional to $(\partial \cdot A)^2$ to the Lagrangian, that it would effectively impose the Lorentz gauge condition. There is a problem in the expression above for the case where $\lambda = 0$. In that case the propagator becomes infinite. Finally, we need to show that physics is indeed insensitive to the choice of λ. To see that, one notices that the field always couples to the current, $j^\alpha(x)A_\alpha(x)$. Once into Fourier space, this will appear as $j^\alpha(k)G_{\alpha\beta}(k)$. The longitudinal piece of the projector, which is proportional to $k^\alpha k^\beta$, will always contract with the current and give a piece $k \cdot j$. This is zero from current conservation. A more sophisticated and involved derivation of how this relates to gauge invariance can be found in Sec. 9.5 of Peskin and Schroeder.

7.8 Massive Vector Field

The Lagrangian for a massive vector field is:

$$L = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2}m^2W^2. \tag{7.41}$$

The corresponding field operators are

$$W^\mu(x) = \int \frac{d^3k}{2E(2\pi)^3} \sum_{s=1,2,3} \epsilon_s^\mu(k) \left\{ e^{-ik\cdot x}a_s(k) + e^{-ik\cdot x}a_s^\dagger(k) \right\}, \tag{7.42}$$

where the three polarization vectors, $\epsilon_{s=1,2,3}$, are purely spatial in the frame of the particle. The creation operators obey the commutation relations,

$$[a_s(k), a_{s'}^\dagger(q)] = 2E(2\pi)^3\delta_{ss'}\delta(k-q). \tag{7.43}$$

One can calculate the propagator,

$$\begin{aligned} G^{\alpha\beta}(x) &= \int \frac{d^3k}{2E_k(2\pi)^3} P^{\alpha\beta}(k)e^{-ik\cdot x}\Big|_{k_0=E_k}, \\ &= \int \frac{d^4k}{k^2 - m^2 + i\epsilon} e^{-ik\cdot x}P^{\alpha\beta}(k). \end{aligned} \tag{7.44}$$

where P is the projection operator that eliminates the vector which is purely timelike in the restframe of the particle,

$$P^{\alpha\beta}(k) = \sum_{s=1,2,3} \epsilon_s^\alpha(k)\epsilon_s^{\dagger\beta}(k) = -g^{\alpha\beta} + \frac{k^\alpha k^\beta}{m^2}. \tag{7.45}$$

As an alternative method for calculating the propagator, one can calculate the Green's function from the equations of motion assuming a coupling $j \cdot W$.

$$\partial^2 W^\rho - \partial^\rho \partial \cdot W + m^2 W^\rho = j^\rho. \tag{7.46}$$

The Green's function is then a solution to the equation,

$$\begin{aligned} K^{\alpha\beta}(k)G_{\beta\gamma}(k) &= g^\alpha_{\ \gamma} \\ K^{\alpha\beta} &= -k^2 g^{\alpha\beta} + k^\alpha k^\beta + m^2 g^{\alpha\beta}. \end{aligned} \tag{7.47}$$

One can solve for G by assuming G is of the form,

$$G^{\alpha\beta} = Ag^{\alpha\beta} + Bk^\alpha k^\beta/k^2, \tag{7.48}$$

where A and B are scalar functions of k, i.e. they depend only on k^2 and m^2. Multiplying KG,

$$K^{\alpha\beta}G_\beta^{\ \gamma} = g^{\alpha\gamma} = -k^2Ag^{\alpha\gamma} + m^2Ag^{\alpha\beta} + Ak^\alpha k^\gamma + m^2Bk^\alpha k^\beta/k^2. \tag{7.49}$$

Solving for A and B,

$$A = \frac{-1}{k^2-m^2},\ B = \frac{k^2/m^2}{k^2-m^2}, \tag{7.50}$$

which is the same propagator as in Eq. (7.45). However, once again the $i\epsilon$s must be added by hand to arrive with the correct time ordering or boundary conditions.

7.9 Spin 1/2 Particles

Beginning with the field operator,

$$\Psi(x) = \sum_s \int \frac{md^3k}{E_k(2\pi)^3}\left\{b_s(k)u_s(k)e^{-ik\cdot x} + d_s^\dagger(k)v_s(k)e^{ik\cdot x}\right\}, \tag{7.51}$$

$$\{b_s(k), b_{s'}^\dagger(q)\} = \frac{E_k}{m}(2\pi)^3\delta_{ss'}\delta(k-q)\cdots,\ \ \bar{u}_s(k)u_{s'}(k) = \delta_{ss'},\ \ \bar{u}_s(k)v_{s'}(k) = 0,$$

it is straightforward to calculate the propagator,

$$\begin{aligned}
S(x) &\equiv -i\mathcal{T}\langle 0|\Psi(x)\bar{\Psi}(0)|0\rangle \qquad (7.52)\\
&= -i\sum_s \int \frac{md^3k}{E_k(2\pi)^3}\left\{u_s(k)\otimes\bar{u}_s(k)e^{-ik\cdot x}\big|_{k_0=E_k}\Theta(x_0) - v_s(k)\otimes\bar{v}_s(k)e^{ik\cdot x}\big|_{k_0=E_k}\Theta(-x_0)\right\}\\
&= -i\int \frac{md^3k}{E_k(2\pi)^3}\left\{\frac{\not{k}+m}{2m}e^{-ik\cdot x}\bigg|_{k_0=E_k}\Theta(x_0) - \frac{-\not{k}+m}{2m}e^{ik\cdot x}\bigg|_{k_0=E_k}\Theta(-x_0)\right\}\\
&= \int \frac{d^4k}{(2\pi)^4}\frac{\not{k}+m}{k^2-m^2+i\epsilon}e^{-ik\cdot x}.
\end{aligned}$$

The second step above used the fact that $u_s\otimes\bar{u}_s$ and $v_s\otimes\bar{v}_s$ were projectors for positive/negative energy solutions, which was derived as a HW problem in chapter 3. The minus sign in the second step comes from the definition of the propagator, which features an extra minus sign for the reverse time-ordering. It is related to the fact that the Fermi fields couple to anti-commuting fields, and that their anti-commutation relations is what is required for the causality relations. The propagator in momentum space is often re-expressed in the manner,

$$S(p) = \frac{\not{p}+m}{p^2-m^2+i\epsilon} = \frac{1}{p^2-m^2+i\epsilon}(\not{p}+m)(\not{p}-m+i\epsilon)(\not{p}-m+i\epsilon)^{-1} = (\not{p}-m+i\epsilon)^{-1}, \tag{7.53}$$

where the "-1" denotes taking the inverse of the matrix. For reasons of notational convenience, this is often written as

$$S(p) = \frac{1}{\not{p}-m+i\epsilon}, \tag{7.54}$$

but can be confusing because the placement of the matrix in the denominator only denotes that it is the inverse matrix.

7.10 Homework Problems

1. For the charged scalar field, the propagator is defined as

$$\begin{aligned} G(x) &= -i\mathcal{T}\langle 0|\Phi(x)\Phi^\dagger(0)|0\rangle, \\ \Phi(x) &= \int \frac{d^3k}{2E(2\pi)^3}\left[a(k)e^{-ik\cdot x} + b^\dagger(k)e^{ik\cdot x}\right], \ [a(k), a^\dagger(q)] = 2E_k(2\pi)^3\delta(k-q). \end{aligned}$$

   Derive the propagator in momentum space,

$$G(k) = \int d^4k e^{ik\cdot x} G(x).$$

2. Beginning with the Lagrangian,

$$L = \frac{-1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m^2W^2, \ \ F^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}\partial_\alpha W_\beta,$$

   find the conjugate field $\pi^\mu(x)$, and show that it commutes with $W(x)$, defined in Eq. (7.42), to give a delta function.

3. Consider the conductivity tensor, defined by

$$J_i = \sigma_{ik}E_k.$$

   (a) Consider the conductivity tensor, defined by

$$J_i = \sigma_{ik}E_k.$$

      i. Show that this can be expressed as

$$\sigma_{ij} = \frac{-i}{\hbar}\int_{-\infty}^{0} dt' \int d^3x' x'_j \ \langle\Psi_0|[J_i(0), \rho(x', t')]|\Psi_0\rangle,$$

      where ρ is the charge density operator and $|\Psi_0\rangle$ is the state in the absence of electric field. Your derivation should begin with the expression in first-order perturbation theory that:

$$|\Psi\rangle = |\Psi_0\rangle - \frac{i}{\hbar}\int_{-\infty}^{0} dt\ V(t)|\Psi_0\rangle.$$

      ii. Show that this can be expressed as:

$$\sigma_{ij} = \frac{i}{\hbar}\int_{-\infty}^{0} dt' \int d^3x' \ t'\langle[J_i(0), J_j(x', t')]\rangle$$

      iii. Consider the 1-D case (no indices) where one defines the quantity,

$$g(t) \equiv \frac{i}{\hbar}\int d^3r \ \langle[J(0), J(r, t)]\rangle.$$

      Show that:

A.

$$g(t) = g^*(t).$$

B.

$$g(t) = -g(-t).$$

C.

$$g^*(\omega) = -g(\omega).$$

D.

$$\sigma = \lim_{\omega \to 0} \frac{\mathcal{I}m\, g(\omega)}{2\omega}.$$

For the Fourier transforms, use the definitions

$$g(t) = \int d\omega e^{-i\omega t} g(\omega)/(2\pi),$$

$$g(\omega) = \int dt e^{i\omega t} g(\omega)$$

8 Perturbation Theory

8.1 Decays in First-Order Perturbation Theory

We consider the decay of a particle to n final-state particles in first-order perturbation theory. The evolution matrix element connecting an initial state $|p_i\rangle$ to a final state $\langle p'_1 \cdots p'_n|$ if the interaction is turned on a long time ($2T$), is:

$$(-i)\int_{-T}^{T} dt d^3r\, \mathcal{T}\langle p'_1 \cdots p'_n|V(\mathrm{r},t)|p_i\rangle. \tag{8.1}$$

The probability of having a transition is:

$$P_{p'_1\cdots p'_n} = \left|\int_{-T}^{T} dt d^3r\, \mathcal{T}\langle p'_1 \cdots p'_n|V(\mathrm{r},t)|p_i\rangle\right|^2, \tag{8.2}$$

$$\tag{8.3}$$

For this example, we consider a particle associated with a scalar field ϕ that is coupled to n other separate scalar fields. The interaction is assumed to have the form,

$$\begin{aligned} V(r,t) &= g\phi(r,t)\phi_1(r,t)\cdots\phi_n(r,t) \qquad (8.4)\\ &= g\int d\tilde{p}d\tilde{p}_1\cdots d\tilde{p}_n\ \left(e^{-ip\cdot x}a(p)+e^{ip\cdot x}a^\dagger(p)\right)\\ &\quad\cdot\left(e^{-ip_1\cdot x}a_1(p_1)+e^{ip_1\cdot x}a_1^\dagger(p_1)\right)\cdots\left(e^{-ip_n\cdot x}a_n(p_n)+e^{ip_n\cdot x}a_n^\dagger(p_n)\right)\end{aligned}$$

where x refers to the four-vector (t,r), g is the coupling constant and $d\tilde{p} = d^3p/[2E_p(2\pi)^3]$. The creation and destruction operators satisfy the relations, $[a(p),a^\dagger(q)] = (2\pi)^3 2E_p\delta^3(p-q)$. Next, we need to express the write the normalized asymptotic states in terms of such creation operators:

$$|p'_1\cdots p'_n\rangle = \prod_{j=1}^{n}\left(\frac{a^\dagger(p'_j)}{\sqrt{2E'_j\Omega}}\right), \tag{8.5}$$

where Ω is the volume. The transition probability is then easy to calculate by inserting the potential into the matrix element,

$$P_{p'_1\cdots p'_n} = \frac{g^2}{\Omega^{n+1}(2E_i)(2E'_1)\cdots(2E'_n)}\left|\int_{-T}^{T} dt_1 d^3r_1\, e^{-i(p_i-p'_1-p'_2\cdots-p'_n)\cdot x_1}\right|^2. \tag{8.6}$$

The integral over coordinate space inside the absolute value sign yields $(2\pi)^3\delta^3(p-p'_1\cdots-p'_n)$. Since the matrix element is squared there are two deleta functions. One delta function can be brought outside, while the other can be replaced by $\delta(0) = \Omega/(2\pi)^3$. This gives

$$\begin{aligned} P_{p'_1\cdots p'_n} &= \frac{g^2(2\pi)^3}{\Omega^n(2E_i)(2E'_1)\cdots(2E'_n)}\delta^3(p-p'_1-\cdots p'_n)\left|\int_{-T}^{T} dt_1\, e^{-i[E_i-(E'_1+E'_2+\cdots E'_n)]t_1}\right|^2 \qquad (8.7)\\ &= 4\frac{g^2(2\pi)^3}{\Omega^n(2E_i)(2E'_1)\cdots(2E'_n)}\delta^3(p-p'_1-\cdots p'_n)\frac{\cos^2\left([E_i-(E'_1+E'_2+\cdots E'_n)]T\right)}{[E_i-(E'_1+E'_2+\cdots E'_n)]^2}.\end{aligned}$$

For decays, we are interested in the rate of decay, as the number should depend on the time $2T$. The rate is

$$\begin{aligned} R_{p'_1\cdots p'_n} &= \frac{1}{2}\frac{dP_{p'_1\cdots p'_n}}{dT} \\ &= 2\frac{g^2(2\pi)^3}{\Omega^n(2E_i)(2E'_1)\cdots(2E'_n)}\delta^3(p-p'_1-\cdots p'_n)\frac{\sin\left(2[E_i-(E'_1+E'_2+\cdots E'_n)]T\right)}{[E_i-(E'_1+E'_2+\cdots E'_n)]} \end{aligned} \tag{8.8}$$

Since $\lim_{\Lambda\to\infty}\sin(\Lambda x)/x = \pi\delta(x)$, the rate becomes:

$$R_{p'_1\cdots p'_n} = \frac{g^2(2\pi)^4}{\Omega^n(2E_i)(2E'_1)\cdots(2E'_n)}\delta^4(p_i-[p'_1+\cdots p'_n]). \tag{8.9}$$

This rate is into specific outgoing states. Summing over all outgoing states involves adding a factor $\prod_j[\Omega d^3p'_j/(2\pi)^3]$. This eliminates the factors of the volume and one finds the decay rate to be

$$\frac{dN}{dt\, d\tilde{p}'_1\cdots d\tilde{p}'_n} = g^2\frac{(2\pi)^4}{2E_i}\delta^4(p_i-[p'_1+\cdots p'_n]). \tag{8.10}$$

EXAMPLE:
Consider a scalar particle of mass M that decays into two equal mass scalar bosons of mass m with decay constant g as described above. Solve for the overall decay rate:

From above (considering the initial particle at rest),

$$\begin{aligned} \Gamma &= g^2(2\pi)^4\frac{1}{2M}\int\frac{d^3p'_1}{(2\pi)^3 2E'_1}\frac{d^3p'_2}{(2\pi)^3 2E'_2}\,\delta(M-2E'_p)\delta^3(p'_1-p'_2) \\ &= g^2\frac{(2\pi)^{-2}}{(2M)(2E'_1)(2E'_2)}\int d^3p'_1\delta\left(M-2\sqrt{m^2+p'^2_1}\right) \\ &= \frac{g^2}{2\pi}\frac{\sqrt{(M/2)^2-m^2}}{4M^2}. \end{aligned} \tag{8.11}$$

It is useful to do dimensional analysis. From the expression for $V(x)$, one can see that g has dimensions of energy. The overall rate then has units of E, which if the $\hbar$s were accounted for, would turn into units of inverse time.

8.2 Cross Sections in First-Order Perturbation Theory

One can calculate rates where the initial state is a two-particle state rather than a one-particle state as considered in the previous subsection. In that case, one can follow the same steps as before. The extra incoming particle introduces an extra factor of $1/(2E_2)$ and an extra factor of $1/\Omega$ from the normalization of the incoming state.

$$\frac{dN}{dt\, d\tilde{p}'_1\cdots d\tilde{p}'_n} = g^2\frac{(2\pi)^4}{(2E_1)(2E_2)\Omega}\delta^4(p_1+p_2-[p'_1+\cdots p'_n]). \tag{8.12}$$

The rate is related to the cross section by the relation, $dN/dt = \sigma v_{\rm rel,cm}/\Omega$, where the relative velocity, $|v_1 - v_2|$ is the value in the center-of-mass frame. This transforms the last expression to:

$$\frac{d\sigma}{d\tilde{p}'_1 \cdots d\tilde{p}'_n} = g^2 \frac{(2\pi)^4}{(2E_{1,\rm cm})(2E_{2,\rm cm})v_{\rm rel,cm}} \delta^4(p_1 + p_2 - [p'_1 + \cdots p'_n]), \tag{8.13}$$

where $E_{1,\rm cm}$ and $E_{2,\rm cm}$ are the energies of the incoming particles in the center of mass frame. As one of the HW problems, you will see that

$$v_{\rm rel,cm} E_{1,\rm cm} E_{2,\rm cm} = \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} \tag{8.14}$$

This gives the final result,

$$\frac{d\sigma}{d\tilde{p}'_1 \cdots d\tilde{p}'_n} = g^2 \frac{(2\pi)^4}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \delta^4(p_1 + p_2 - [p'_1 + \cdots p'_n]), \tag{8.15}$$

8.3 Problems in higher-order Perturbation theory

Before listing the rules for Feynmann diagrams for arbitrary order, we consider a simple second-order process. In this process we consider the scattering of two particles of type a with p_1 and p_2 to final states of particles of type b with momenta p'_1 and p'_2. We will consider an interaction element of the form,

$$V(x) = \frac{g}{2}\Phi_a^2(x)Z(x) + \frac{g}{2}\Phi_b^2(x)Z(x), \tag{8.16}$$

where Z is also a real scalar field with masses M_Z, while the masses of the Φ_a and Φ_b fields will be denoted by m_a and m_b. The transition element will be a two-step process. After performing the contractions of field operators with the incoming and outgoing states, the matrix element differs from the one step-case considered earlier,

$$\begin{aligned}
& (-i)g \int_{-T}^{T} dt_1 d^3x_1 \, e^{-i(p_1+p_2-p'_1-p'_2)\cdot x_1} \rightarrow \\
& (-i)^2 g^2 \int_{-T}^{T} dt_1 d^3x_1 \, e^{-i(p_1+p_2)\cdot x_1} \int_{-T}^{T} dt_2 d^3x_2 e^{i(p'_1+p'_2)\cdot x_2} [-iG_F(x_1 - x_2)] \\
= \; & (-i)^2 g^2 \int_{-T}^{T} dt_1 d^3x_1 \, e^{-i(p_1+p_2-p'_1-p'_2)\cdot x_1} \int_{-T}^{T} d^4(x_1 - x_2) e^{i(p'_1+p'_2)\cdot(x_2-x_1)} [-iG_F(x_1 - x_2)] \\
= \; & (-i)^2 g^2 \int_{-T}^{T} dt_1 d^3x_1 \, e^{-i(p_1+p_2-p'_1-p'_2)\cdot x_1} [-iG_F(p'_1 + p'_2)].
\end{aligned} \tag{8.17}$$

First some comments about the factor of $1/2$ in the expression above for V. Since either incoming particle could have contracted with a given factor of Φ in $V(x_1)$, it eliminated the factor of 2. Secondly, the usual factor of $1/n!$ in the expression for the perturbation theory, $[(-i)^2/n!]\mathcal{T}\langle V(x_1)\cdots V(x_n)\rangle$, disappears because the incoming state could have coupled to either $V(x_1)$ or $V(x_2)$. The factor $(-i)$ preceeding G_F derives from the i in the definition of G_F in Eq. (7.2).

By inspection, one can now see that the results from the previous section will follow over exactly, except that the coupling constant will be replaced by:

$$-ig \rightarrow (-ig)^2[-iG_F(p_1' + p_2')]. \tag{8.18}$$

Due to momentum conservation, $p_1' + p_2'$ could equally well have been expressed as $p_1 + p_2$. The process above is well described pictorially through "Feynmann diagrams":

The rules for calculating Feynmann diagrams for cross sections are as follows:

1. Draw all the diagrams to the order you wish to calculate. Each diagram will represent a different contribution to the overall amplitude. The various contributions will add coherently. For the example above, there is only one diagram.
2. For each internal line, associate a momentum as determined by four-momentum conservation. For this example, the momentum is $(2E_i, 0, 0, 0)$.
3. Express the cross section as:

$$\frac{d\sigma}{d\tilde{p}_1' \cdots d\tilde{p}_n'} = \frac{(2\pi)^4 \delta(p_1 + p_2 - [p_1' + \cdots p_n'])}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \left|\mathcal{T}(p_1, p_2 \rightarrow p_1' \cdots p_n')\right|^2. \tag{8.19}$$

4. Calculate $\mathcal{T}$ by making a product of $(-i)$ multiplied by the coupling constants for each vertex, multiplied by the propagator $-iG_F(k)$ for each internal line. Note that $\mathcal{T}$ is referred to as the scattering amplitude, and is related to something known as the "T-matrix" in scattering theory. Try not to get it confused with the time-ordering operator.

The rules get somewhat more complicated for Fermions or for vector particles. For decay rates, the procedure is the same, except the factor $4\sqrt{\{\cdots\}}$ in the denominator is replaced by $2\times$ the invariant mass of the decaying particles.

EXAMPLE:
Calculate the total cross section from the example discussed above.

The cross section is

$$\begin{aligned}
\sigma &= \int \frac{d^3p_1' d^3p_2'}{(2\pi)^6 (2E_1')(2E_2')} \frac{(2\pi)^4 \delta^4(p_1 + p_2 - p_1' - p_2')}{4\sqrt{(p_1 \cdot p_2)^2 - m_a^4}} \left|\mathcal{T}\right|^2, \\
\mathcal{T} &= \frac{ig^2}{K^2 - M_Z^2 + i\epsilon} = \frac{-ig^2}{s - M^2}.
\end{aligned}$$

Exploiting the spatial aspects of the delta function,

$$\begin{aligned}
\sigma &= \frac{|\mathcal{T}|^2}{(2\pi)^2 4E_a^2 \sqrt{[2E_1^2 - m_a^2]^2 - m_a^4}} \cdot 4\pi \int p'^2 dp' \, \delta(2E_p' - \sqrt{s}) \\
&= \frac{|\mathcal{T}|^2}{2\pi s} \frac{\sqrt{s - 4m_b^2}}{\sqrt{s - 4m_a^2}}.
\end{aligned}$$

8.4 Fermions and Derivative Couplings

For fermions, the normalization of the wave function has a factor E/m as opposed to the factor $2E$ for bosons. Thus, the external lines for fermions are assigned a factor m/E rather than $1/2E$. A second difference is that the the vertices pick up a factor of $u_{s,i}(p)$, $\bar{u}_{s,i}(p)$, $v_{s,i}(p)$ or $\bar{v}_{s,i}(p)$. This comes from that fact that $\langle 0|\Psi_i(x)b_s^\dagger(p|0\rangle = u_{s,i}(p)e^{-ip\cdot x}$. For a vertex connected to an incoming line, one adds a term $u_{s,i}(p)$ if the incoming line is a particle and a term $\bar{v}_{s,i}(p)$ if an anti-particle is incoming. For the outgoing particles, the vertices pickup the corresponding terms, $\bar{u}_{s,i}(p)$ and $v_{s,i}(p)$. Since fermions carry some sort of conserved charge, arrows are often used to denote the direction of current. For electrons, the arrows usually go with the momentum, while for positrons the arrows go in the opposite direction of the current. Although we won't talk about loops in this chapter, we mention that fermion loops give an extra minus sign.
Most cross sections involve summing over polarizations. When that is the case it is often convenient to use projectors to calculate the squared matrix element. For instance, if one of the incoming lines is an electron with momentum p, one would insert $u_{s,i}(p)$ into the vertex connecting the incoming line. Further, when squaring the matrix element, one would encounter $\bar{u}_{s,j}(p)$ in the complex conjugate of the matrix element. The product of the two, $u(p)_{s,i}\bar{u}(p)_{s,j}$, summed over spins s, is the projector, $(\not{p}+m)/2$. In this way, the spins are no longer specifically mentioned as one would expect given they are summed over. Furthermore, the Dirac indices are also contracted with the indices of the scattering element. This means that squared matrix elements can be written as traces over products of Dirac matrices (such as $\not{p}$). An example of this is given further below.
Some interaction terms involve derivatives. Examples are:

- pion, nucleon: $g_{\pi NN}(\partial_\mu\vec{\pi})\cdot\bar{\Psi}\vec{\tau}\gamma^\mu\gamma^5\Psi$.
- photon to charged scalar field, $eA_\mu\Phi^*(i\partial^\mu\Phi) - eA_\mu(i\partial^\mu\Phi^*)\Phi$.

If the derivatives act on fields which are contracted internally (they are part of an internal propagator), the derivatives can be replaced with the momenta involved in the relevant propagators, $\partial_\mu \to -ik_\mu$. If the derivatives act on fields that are contracted with the initial or final state, they can be replaced with the momenta of the incoming or final states.

EXAMPLE:
Calculate the differential cross section for the process of a charged scalar of mass M annihilating with its antiparticle electromagnetically, then creating an electron positron pair, where the electron has mass m. Let the incoming momenta be p_1 and p_2, while the outgoing momenta are p_1' and p_2'.

The scattering amplitude can be inferred from the following diagram:

If you multiply the diagram times it complex conjugate,

one finds the squared matrix element, which has products of $u_s\bar{u}_s$ and $v_s\bar{v}_s$. If one sums over the spins, the combinations $\sum_s u_{s,i}(p)\bar{u}_{s,j}(p)$ are projectors and can be written as $(\not{p}+m)_{ij}$. Similarly, the combinations with vs can be written as $(-\not{p}+m)_{ij}$. The squared amplitude can then be expressed as:

Here, the cuts refer to the fact that the two internal fermion propagators are on-shell, i.e., they are replaced by $(\not{p}+m)/2m$ if the momentum is defined in the same di-

rection as the arrows, and $(-\not{p}+m)/2m$ if the momentum p is going opposite the direction of the arrows. Drawing the dashed line is known as "cutting" the diagram.

The cross section can now be written as:

$$\begin{aligned} d\sigma &= \frac{(2\pi)^4}{4\sqrt{(p_1\cdot p_2)^2-M^4}}\,|\mathcal{T}|^2\,\delta^4(p_1+p_2-p_1'-p_2')d\tilde{p}_1'd\tilde{p}_2', \qquad (8.20)\\ |\mathcal{T}|^2 &= e^4(-i)(p_1-p_2)_\mu(-i)G^{\mu\nu}(p_1+p_2)\Pi_{\nu\alpha}iG^{\alpha\beta}(p_1+p_2)i(p_1-p_2)_\beta,\\ \Pi_{\nu\alpha} &\equiv \mathrm{Tr}\,\gamma_\nu(\not{p}_1'+m)\gamma_\alpha(-\not{p}_2'+m). \end{aligned}$$

Again, $d\tilde{p} = d^3p/[2E_p(2\pi)^3]$, even though the normalization for fermions had a factor of m/E_p rather than $1/2E_p$. However, absorbing the factors of $1/2m$ from the projectors $(\not{p}+m)/2m$ accounted for the factor. The fermion loop in the middle of the diagram has all its Dirac indices summed over, hence it can be expressed as a trace. The vertices associated with the incoming particles bring in factors of $e(p_1-p_2)^\mu$ from the coupling of the charged scalar field to the electromagnetic field, $e[(-i\partial_\mu\Phi^*)\Phi+\Phi^*(i\partial_\mu\Phi)]A^\mu$.

Calculating the trace in the definition of Π is not too difficult. First, it is easy to see that any trace of an odd number of gamma matrices is zero. Secondly, since $\{\gamma^\alpha,\gamma^\beta\}=g^{\alpha\beta}$, it is easy to see that

$$\mathrm{Tr}\,\gamma^\alpha\gamma^\beta = 4g^{\alpha\beta}, \qquad (8.21)$$

with the factor of four coming from the four Dirac indices. For traces of a larger number of Dirac matrices, it is clear that one always needs an even number of matrices with each Lorentz index if the trace is to be zero, e.g., $\mathrm{Tr}\,\gamma^0\gamma^x\gamma^y\gamma^0=0$. Using the anti-commutation rules, one can readily convince oneself that

$$\mathrm{Tr}\gamma^{\alpha_1}\cdots\gamma^{\alpha_n} = 4\sum_{\text{pair arrangements}}\prod_{ij} g^{\alpha_i\alpha_j}(-1)^P. \qquad (8.22)$$

The sum is over all ways to pair up the n matrices into pairs ij, and P is the number of permutations involved in rearranging the matrices so that they would be paired with neighbors. For instance,

$$\mathrm{Tr}\,\gamma^\alpha\gamma^\beta\gamma^\delta\gamma^\eta = 4\left[g^{\alpha\beta}g^{\delta\eta}-g^{\alpha\delta}g^{\beta\eta}+g^{\alpha\eta}g^{\beta\delta}\right]. \qquad (8.23)$$

This means that

$$\mathrm{Tr}\,\not{p}_1'\gamma^\alpha\not{p}_2'\gamma^\beta = 4\left[2p_1'^\alpha p_2'^\beta - g^{\alpha\beta}p_1'\cdot p_2'\right], \qquad (8.24)$$

and

$$\Pi^{\nu\alpha} = 4\left[-p_1'^\nu p_2'^\alpha - p_2'^\nu p_1'^\alpha + p_1'\cdot p_2' g^{\nu\alpha} + m^2 g^{\nu\alpha}\right]. \qquad (8.25)$$

Finally, we are in a position to begin calculating. We will let the two incoming particles have momentum $p\hat{z}$ and $-p\hat{z}$, which gives $p_1-p_2=(0,0,0,2p)$. For the photon propagator (working in the Feynmann gauge) the only component we then

need to worry about is $G_{zz} = i/([p_1 + p_2]^2 + i\epsilon) = i/s$. This also means that we only need worry about Π_{zz},

$$\Pi_{zz} = 4\left[2p_z'^2 + E'^2 + p'^2 - m^2\right] = 8p'^2(1+\cos^2\theta), \tag{8.26}$$

where θ is the scattering angle, and $\vec{p}' \equiv \vec{p}_1' = -\vec{p}_2'$. The square root in the denominator of Eq. (8.20) is

$$\sqrt{(p_1\cdot p_2)^2 - M^4} = \sqrt{(E_p^2 + |\vec{p}|^2)^2 - M^4} = \sqrt{(2|\vec{p}|^2 + M^2)^2 - M^4} = 2|\vec{p}|E_p. \tag{8.27}$$

The differential cross section is then

$$\begin{aligned} d\sigma &= e^4 \frac{2\pi\delta(2E'-2E)}{8Ep}\frac{d^3p'}{4E'^2(2\pi)^3}\frac{4p^2}{s^2}8p'^2(1+\cos^2\theta), \qquad (8.28)\\ \frac{d\sigma}{d\Omega} &= \frac{e^4pp'^3}{2\pi^2 s^3}(1+\cos^2\theta). \end{aligned}$$

8.5 Connected and Disconnected Diagrams

Thus far, we have only considered Feynmann diagrams where every line and vertex is involved in connecting the initial to the final state. However, there could easily be diagrams that involve fluctuations from the vacuum. For instance, part of the interaction $\bar{\Psi}\gamma^\mu\Psi A_\mu$ describes the creation of a photon along with a particle-antiparticle pair. A Second interaction at a later time could describe the annihilation of the same particles. Diagrammatically, this would be
Any process, from scattering to decays to simply the vacuum evolving to the vacuum, would include arbitrary numbers of disconnected diagrams. Fortunately, as we will show below, these diagrams only contribute to an overall phase of the evolution matrix and can be safely neglected. To see this consider all diagrams of order N contributing to the evolution of the vacuum to itself. Such diagrams arise from the terms in the perturbative expansion,

$$\frac{(-i)^N}{N!}\mathcal{T}\langle \text{out}|V(x_1)\cdots V(x_N)|\text{in}\rangle. \tag{8.29}$$

Once one has contracted all the fields in all possible ways, one can write the term above as

$$\frac{1}{N!}\sum_{\langle n_k\rangle,\ \text{s.t.}\ \sum kn_k = N}\frac{N!}{(\prod_i (in_i)!)}\prod_i (C_i)^{n_i}\frac{(in_i)!}{n_i!}. \tag{8.30}$$

Here, C_k is the sum of all connected diagrams of order k. For any such diagram there are $k!$ ways of arranging the vertices $1\cdots k$, and C_k references only one such arrangement, hence there is no additional $1/k!$. The permutation count in the middle of the expression, $N!/\prod(kn_k)!$ counts the number of ways to assign the N labels, $x_1\cdots x_n$ into connected diagrams of a certain order. Once one has (kn_k) labels assigned to the connected diagrams of order k, there are $(kn_k)!/n_k!$ different ways to assign the kn_k labels to the n_k diagrams of order k. Again, we point out

that since C_k refers to one specfic arrangement of the labels within a closed diagram, and since all such $i!$ arrangements give the same contribution, one should not include the $1/k!$ term in the counting. Finally, after canceling all the factors that appear identically in numerators and denominators, one finds

$$\begin{aligned}\sum_N \frac{1}{N!}\mathcal{T}\langle \text{out}|V(x_1)\cdots V(x_N)|\text{in}\rangle. &= \sum_N \sum_{\langle n_k\rangle,\ \text{s.t.}\ \sum k n_k = N} \prod_k \frac{C_k^{n_k}}{n_k!} \\ &= \sum_{n_1,n_2\cdots} \prod_i \frac{C_k^{n_k}}{n_k!} \\ &= \exp\left\{\sum_i C_k\right\}.\end{aligned} \tag{8.31}$$

The C_ks are all imaginary, i.e., $\sum_k C_k = i\phi$, and the disconnected diagrams summed to all orders provide simply a phase factor. Thus, they can be ignored. If one were to do the same for some other evolution element (besides vacuum-to-vacuum) one would find the same result, that the disconnected pieces can be ignored. For what it's worth, if calculated, the disconnected pieces sum to an amount proportional to the time for which the interaction is turned on. Dividing by this time gives the correction to the vacuum energy due to interactions.

8.6 Homework Problems

1. Derive Eq. (8.14). Begin with the definitions,

$$v^\mu_{\text{cm}} = \frac{p_1'^\mu}{E_{1,\text{cm}}} - \frac{p_2'^\mu}{E_{2,\text{cm}}}, \quad p_i'^\mu \equiv p_i^\mu - P^\mu \frac{P\cdot p_i}{P^2}, \quad E_{\text{i,cm}} = \frac{P\cdot p_i}{\sqrt{P^2}}, \quad P^\mu \equiv p_1^\mu + p_2^\mu.$$

2. For the expression above, consider two particles of the same mass m approaching one another with momenta $p\hat{z}$ and $-p\hat{z}$. Using the result from the previous problem, derive all four components of v^μ_{cm} in terms of p and m.

3. Calculate $\text{Tr}\ \gamma^\alpha \not{p}\gamma^\beta \not{q}\gamma^\delta \not{k}$.

9 Loops

9.1 Loops in Vacuum Diagrams

Loops suck. Any time a loop appears in a Feynmann diagram, there are internal propagators for which the momenta are not determined by the asymptotic momenta. This means that the Feynmann diagrams include integrations over momenta, which adds a term $\int \frac{d^4k}{(2\pi)^4}$ to the evaluation. For instance consider the amplitude for photon-photon scattering:
There is an additional free momentum k. All four components of k are free to take on any value. The amplitude for this particular diagram then becomes:

$$W^{\alpha\beta\delta\eta} = \int \frac{d^4k}{(2\pi)^4} \cdot \mathrm{Tr}\, \gamma^\alpha \frac{1}{\not k - m + i\epsilon}\gamma^\beta \frac{1}{\not k + \not p_2 - m + i\epsilon}\gamma^\delta \frac{1}{\not k - \not p_1 - \not p_1' - m + i\epsilon}\gamma^\eta \frac{1}{\not k - \not p_1 - m + i\epsilon}. \tag{9.1}$$

The difficulty posed by the integral is that it is logarithmically divergent as $k \to \infty$. The procedure for eliminating the most divergent pieces is known as "renormalization theory", though it is more of a scheme than a theory. The divergence points to a sickness in the theory that can be "fixed" by adding counter terms to the Lagrangian that conveniently cancel the offending divergence. We will put off such calculations for the next course. In condensed matter calculations of infinities can be canceled by cutoffs in k, which might be physical, e.g., the Debye frequency. However, in particle physics the theory simply falls apart below a certain length scale, even after renormalization.
Loops with fermions are more divergent than loops with bosons since the propagators scale as $1/k$ rather than $1/k^2$. Loops with more vertices have fewer problems since each propagator in the loop brings along a factor of $1/k$. Adding derivatives to the couplings makes things worse. Fortunately, for electromagnetism the coupling is $\bar\Psi\gamma^\mu\Psi A_\mu$ and has no derivatives. It can be shown that for theories with coupling constants that are either dimensionless or have dimensions of mass, that the number of divergent diagrams is finite.
The most divergent diagrams are those with a minimum number of vertices. These would be the self-energy diagrams that represent the change in the energy for a mode of a specific wave number. Self-energy diagrams refer to diagrams with one line in and one out. They refer to changes to a particle's energy/momentum relation (a.k.a. its dispersion relation) due to interaction. Decay rates are also related to the self energy. The self-energy also describes charge renormalization, and in QCD is used to describe the famous running coupling constant. We confine our discussion to a few simple examples, with the purpose of explaining a few principles, but without delving into renormalization the scope of this chapter will be rather limited.
We begin by considering a real scalar field that is coupled to a fermion field,

$$V(x) = g\Phi(x)\bar\Psi(x)\Psi(x). \tag{9.2}$$

We then consider the propagation of the scalar particle to itself. The lowest-order correction to the propagator is described by the diagram:

Including the effect of the Propagator to this order,

$$-iG(k) = -iG_0(k) - iG_0(k)(-i\Pi(k))(-i)G_0(k), \; -i\Pi(k) = g^2 \int \frac{d^4q}{(2\pi)^4} \mathrm{Tr} \frac{1}{\not{q} - m + i\epsilon} \frac{1}{\not{q} - \not{k} - m} \tag{9.3}$$

Note that the propagator is $-iG$. The choice of defining the loop as $-i\Pi$ rather than as Π will be clear later. The difficulty with the expression for $\Pi(k)$ is that it is divergent for large k. To rectify this, one can implement an upper cutoff in momentum, Λ. Dimensionally, $\Pi(k)$ then has a contribution proportional to Λ^2. To eliminate this divergence, one can assume the Lagrangian also had a mass term, with $M^2 \propto \Lambda^2$ that fortuitously cancels the divergence. There are a couple of schemes for handling such divergences, most famously dimensional regularization. The purpose of these schemes is to make $\Pi(k)$ finite.
Surprisingly, it is easy to simultaneously incorporate all diagrams of the form,
Algebraically,

$$\begin{aligned} -iG(k) &= -iG_0(k) - iG_0(k)\Pi(k)(-i)G_0(k) - iG_0(k)\Pi(k)(-i)G_0(k)\Pi(k)(-i)G_0(k) + \quad (9.4) \\ &= -iG_0(k) - iG_0(-i\Pi(k))(-i)G(k) \\ -iG(k) &= \frac{-iG_0(k)}{1 - G_0(k)\Pi(k)}. \end{aligned}$$

Since $G_0 = -1/(k^2 - m^2 + i\epsilon)$, one finds

$$-iG(k) = \frac{i}{k^2 - m^2 - \Pi(k) + i\epsilon}. \tag{9.5}$$

The self-energy Π has units of energy squared, and alters the positions of the poles. If Π is independent of k, it corresponds to an additional contribution to the mass, but if Π has a k dependence one must solve for the new strengths to the poles. In the vacuum Lorentz invariance demands that $\Pi(k)$ is of the form $\alpha k^2 + M_\Pi^2$. The part proportional to k^2 alters the strength of the poles. When doing the analogous calculation for photon fields, the change in the strength of the poles is absorbed into the strength of the coupling to the photon. The mass correction can simply be ignored by redefining the "bare" mass of the particle.
For the remainder of this subsection, we will ignore the real part of Π and consider the imaginary part. An imaginary part will be non-zero only if there are ways to cut the self-energy diagram and have all the lines be on shell. Considering only the imaginary part of Π and assuming it is small, one can write the propagator as

$$-iG(k) = \frac{i}{(k_0 - E_k - i\Pi_I/2E_k)(k_0 + E_k + i\Pi_I/2E_k)}. \tag{9.6}$$

This means that the position of the poles has moved off the real axis by an amount $-\Pi_I/E_k$. In a wavefunction, the phase picks up a factor of $e^{iEt - \Gamma t/2}$, where the decay rate for a particle at rest is

$$\Gamma = |\Pi_I|/M. \tag{9.7}$$

This simple expression only applies for particles decaying with energies close to the on-shell energy. Once one considers far off-shell particles, one needs to be more careful. The range of

energies which one needs to consider also depends on $\mathbf{\Pi_I}$. For instance, if $\mathbf{\Pi_I} \to \mathbf{0^-}$, then it only affects the propagator as you approach being on-shell, and when that neighborhood is small one can neglect the $\boldsymbol{k}$ dependence of $\mathbf{\Pi_I}$.

EXAMPLE:
Calculate the decay rate of a real scalar particle of mass $\boldsymbol{M}$ that decays to a fermion-antifermion pair, where the masses of the fermions are less than half $\boldsymbol{M}$, and the coupling constant $\boldsymbol{g}$ is small.

The self energy is

$$\begin{aligned}
\mathbf{\Pi_I}(\boldsymbol{k}) &= \mathbf{Re}\, \boldsymbol{g^2} \int \frac{\boldsymbol{d^4q}}{\boldsymbol{(2\pi)^4}} \mathbf{Tr}\, \frac{\boldsymbol{i}}{(\not{q} - \boldsymbol{m} + \boldsymbol{i\epsilon})} \frac{\boldsymbol{i}}{(\not{q} + \not{k} - \boldsymbol{m} + \boldsymbol{i\epsilon})} \qquad (9.8)\\
&= \mathbf{Re}\, \boldsymbol{g^2} \int \frac{\boldsymbol{d^4q}}{\boldsymbol{(2\pi)^4}} \mathbf{Tr}\left[(\not{q} + \boldsymbol{m})(\not{q} - \not{k} + \boldsymbol{m})\right] \frac{\boldsymbol{i}}{[\boldsymbol{q^2} - \boldsymbol{m^2} + \boldsymbol{i\epsilon}]} \frac{\boldsymbol{i}}{[\boldsymbol{(k-q)^2} - \boldsymbol{m^2} + \boldsymbol{i\epsilon}]}\\
&= \frac{\boldsymbol{g^2}}{\boldsymbol{(2\pi)^4}} \int \frac{\boldsymbol{d^3q}}{\boldsymbol{2E_q}} \boldsymbol{(2\pi^2)\delta((k_0 - q_0)^2 - E^2_{k-q})} \mathbf{Tr}\left[(\not{q} + \boldsymbol{m})(\not{q} - \not{k} + \boldsymbol{m})\right]
\end{aligned}$$

To perform the last step, one needed to first perform the integration over $\boldsymbol{q_0}$, which picks up two poles (there are 4 over all). We write only the term for one pole, then multiply by two as the two contributions should be equal. Then, the identity $\boldsymbol{1/(x + i\epsilon) = \mathcal{P}(1/x) + \pi i \delta(x)}$ was used to separate the imaginary part, which is what is needed to arrive at the real part given that each pole brought out a factor of $\boldsymbol{2\pi i}$ when integrating over $\boldsymbol{q_0}$. The trace is easily calculated,

$$\begin{aligned}
\mathbf{Tr}\,(\not{q} + \boldsymbol{m})(\not{q} - \not{k} + \boldsymbol{m}) &= \boldsymbol{4\left[q^2 + m^2 - k \cdot q\right]} \qquad (9.9)\\
&= \boldsymbol{-4(ME_q - 2m^2) = -8|q|^2.}
\end{aligned}$$

In the last step, it was assumed that $\boldsymbol{k = (M, 0, 0, 0)}$, and since the fermions are on-shell, $\boldsymbol{q^2 = m^2}$ with $\boldsymbol{E_q = M/2}$. Finally, this yields

$$\begin{aligned}
|\mathbf{Im}\, \mathbf{\Pi}| &= \frac{\boldsymbol{g^2}}{\boldsymbol{8M\pi^2}} \int \boldsymbol{4\pi q^4 dq\, \delta((k_0 - E_q)^2 - E_q^2)} \qquad (9.10)\\
&= \frac{\boldsymbol{g^2}}{\boldsymbol{\pi}} \frac{\boldsymbol{|q|^3}}{\boldsymbol{M}},
\end{aligned}$$

where $\boldsymbol{q = \sqrt{(M/2)^2 - m^2}}$ is the magnitude of the momentum of an outgoing fermion. The full decay width is then

$$\mathbf{\Gamma} = \frac{\boldsymbol{g^2}}{\boldsymbol{\pi}} \frac{\boldsymbol{|q|^3}}{\boldsymbol{M^2}}. \qquad (9.11)$$

The width going to zero as $\boldsymbol{q \to 0}$ comes from the lack of phase space for decays. If one had solved the analogous problem for the decay of a vector particle, such as a $\boldsymbol{\rho}$ meson, additional powers of $\boldsymbol{q}$ would have shown up due to the vector coupling.

9.2 Propagators in Matter

When evaluating propagators, we throw away terms of the form $\mathcal{T}\langle 0|a^\dagger(t_1)a(t_2<t_1)|0\rangle$ due to the fact that the destruction operator annihilated the vacuum. This changes once the vacuum is replaced by a medium with non-zero occupation numbers, such as for the electrons in a metal. As an example, we consider the propagator for neutrons in a neutron star. The propagator is

$$\begin{aligned}-iS_F(k) = \int d^x k e^{ik\cdot x}\mathcal{T}\langle\Psi(0)\Psi(x)\rangle &= \int d^4x e^{ik\cdot x}\sum_{s,s'}\int \frac{md^3p}{(2\pi)^3E_p}\frac{md^3p'}{(2\pi)^3E'_p} \\ \cdot\Big\{\Big\langle[b_s^\dagger(p)\bar u_s(p)+d_s(p)\bar v_s(p)][b_{s'}(p')u_{s'}(p')e^{-ip'\cdot x}+d_{s'}^\dagger(p')v_{s'}(p')e^{ip'\cdot x}]\Big\rangle\Theta(x_0) \\ -\Big\langle[b_{s'}(p')u_{s'}(p')e^{-ip'\cdot x}+d_{s'}^\dagger(p')v_{s'}(p')e^{ip'\cdot x}][b_s^\dagger(p)\bar u_s(p)+d_s(p)\bar v_s(p)]\Big\rangle\Theta(-x_0)\Big\}.\end{aligned} \tag{9.12}$$

For all the terms involving the anti-particle creation and destruction operators, nothing changes when the initial state is changed from the vacuum to those particles. The only terms that change are those involving the bs and $b^\dagger$s where the energies are below the Fermi energy. The terms with ordered $bb^\dagger$ must then be replaced with those ordered as $b^\dagger b$. The propagator is then equal to the vacuum propagator plus the difference multiplied by the chance the state is occupied, $f_s(p)$.

$$\begin{aligned}-iS_F(k) &= -iS_F(k)_{\text{vacuum}} \\ &+ \int d^4x e^{ik\cdot x}\sum_s\int\frac{md^3p}{(2\pi)^3E_p}f_s(p) \\ &\cdot\left\{\bar u_s(p)u_s(p)e^{-ip\cdot x}\Theta(x_0)+u_s(p)e^{-ip\cdot x}\bar u_s(p)\Theta(-x_0)\right\}.\end{aligned} \tag{9.13}$$

Since $\sum_s u_s(p)\bar u(p)$ is the projector $(\not p+m)/2m$, one can see that the propagator becomes

$$-iS_F(k) = \frac{i(\not k+m)}{k^2-m^2+i\epsilon}+f_s(k)\frac{(\not k+m)}{2E_k}2\pi\delta(k_0-E_k). \tag{9.14}$$

The most common use of the above relation is to calculate the correction to the self energy of a particle moving through matter. For example, one can consider the case of an electron neutrino moving through a gas of electrons, e.g. through the sun. In that case the diagrams of interest are where one need only consider the correction to the electron propagator. The correction to the fermion self energy due to the W is:

$$\begin{aligned}\Sigma_W(k) &= \int\frac{d^3p}{(2\pi)^3}f(p)\left(\frac{g}{2\sqrt{2}}\gamma^\mu(1-\gamma_5)\right)\left[\frac{\not p+m}{2E_p}\right]\left(\frac{g}{2\sqrt{2}}\gamma^\nu(1-\gamma_5)\right) \\ &\cdot\frac{g_{\mu\nu}-(p-k)_\mu(p-k)_\nu/M_W^2}{(p-k)^2-M_W^2}.\end{aligned} \tag{9.15}$$

The correction due to the Z (the tadpole diagram) is:

$$\begin{aligned}\Sigma_Z(k) &= -\frac{g}{4\cos\theta_w}\gamma^\mu(1-\gamma_5)\frac{g_{\mu\nu}-q_\mu q_\nu/M_Z^2}{q^2-M_Z^2} \\ &\cdot\mathrm{Tr}\left\{\int\frac{d^3p}{(2\pi)^3}f(p)\left(\frac{g}{4\cos\theta_w}\gamma^\nu(1-4\sin^2\theta_w-\gamma_5)\right)\left[\frac{\not p+m}{2E_p}\right]\right\},\end{aligned} \tag{9.16}$$

where $q = 0$. Note that electroweak vertices are listed at the end of this chapter.
Note that $\Sigma(k) = \Sigma_Z(k) + \Sigma_W(k)$ is a four-by-four matrix in terms of Dirac indices. The neutrino's propagator becomes

$$-iS_F(k) = \frac{i}{\not{k} - m - \Sigma(k) + i\epsilon}. \tag{9.17}$$

Due to the fact that Σ is a matrix, solving for the dispersion relation, i.e., the energy k_0 for a given mode k, is not so straight-forward. Before proceeding, we make the approximations that the W mass is much larger than any other scale in the problem, and that the electron's are non-relativistic. Further, one can use the fact that $f(p) = f(-p)$, to write a simplified version of the self-energy,

$$\begin{aligned}
\Sigma_W(k) &= -n_e \frac{g^2}{32 M_W^2} \gamma^\mu (1-\gamma_5)(\gamma_0 + 1)\gamma_\mu (1-\gamma_5), \qquad (9.18)\\
&= -n_e \frac{g^2}{16 M_W^2} \gamma^\mu (1-\gamma_5)\gamma_0 \gamma_\mu \\
&= n_e \frac{g^2}{4 M_W^2} \gamma_0 (1-\gamma_5). \\
\Sigma_Z(k) &= n_e \frac{g^2}{16 M_Z^2 \cos^2\theta_w} \gamma_0 (1-\gamma_5)(1 - 4\sin^2\theta_w) \qquad (9.19)
\end{aligned}$$

where $n_e = 2\int f(p) d^3p/(2\pi)^3$ is the density of electrons. In the intermediate steps, the relation that $(1+\gamma_5)(1-\gamma_5) = 0$ was used along with the fact that γ_5 anti-commutes with the γ matrices. To calculate the eigenstates, one needs to find the new values of p_0 for which one has poles in the propagator,

$$(\not{p} - m - \Sigma_Z - \Sigma_W)\psi = 0. \tag{9.20}$$

To simplify, we rewrite the self-energies,

$$\Sigma_Z + \Sigma_W = A\gamma_0(1-\gamma_5), \tag{9.21}$$

where A can be read off Eq.s (9.18). Solving for p_0,

$$\begin{aligned}
(p_0 - A)\psi &= (p\cdot\alpha + m\gamma_0 - A\gamma_5)\psi \qquad (9.22)\\
(p_0 - A)^2 &= \mathrm{p}^2 + m^2 + A^2 - 2A\mathrm{p}\cdot\alpha\gamma_5,
\end{aligned}$$

where we have made use of the fact the $\gamma_5^2 = 1$ and that γ_5 anti-commutes with γ_0, but commutes with α. Moving all the terms proportional to the unit matrix to the l.h.s. and squaring,

$$\begin{aligned}
\left[(p_0 - A)^2 - \mathrm{p}^2 - m^2 - A^2\right]^2 &= 4A^2p^2, \qquad (9.23)\\
(p_0 - A)^2 - \mathrm{p}^2 - m^2 - A^2 &= \pm 2A|p|.
\end{aligned}$$

Thus, for each p there are two energies, due to the fact that the self energies behave like potentials that interact with the spin through the $(1-\gamma_5)$ terms. Next, one assumes that the neutrino energies are much larger than either A or the mass. If one defines $E_p = \sqrt{p^2 + m^2}$, and $p_0 = E_p + \Delta E$, one can solve for ΔE,

$$\Delta E = A \pm A. \tag{9.24}$$

For a given $\boldsymbol{p}$, one state is basically unchanged by the medium, while the other is shifted by $\mathbf{2}\boldsymbol{A}$. Note that the shift of the energy is independent of the momentum.
For neutrinos moving through the sun, there are analogous contributions to the self energy arising from the coupling to the quarks inside the protons. Aside from some factors of $\mathbf{sin^2}\,\boldsymbol{\theta_W}$, these contributions are the same. MSW neutrino oscillations arise from mixing terms that change electron neutrinos to either tau or muon neutrinos. These small mixings are amplified when the energies of the different neutrinos become equal for identical momentum. This happens when the energy shifts listed above cancel the energy difference due to differences in neutrino masses. Since the masses are much smaller than the neutrino energies, one can state that

$$\boldsymbol{E_p \approx |p| + \frac{m^2}{2|p|}}. \tag{9.25}$$

For $\boldsymbol{\Delta E}$ to push you into a condition where there is maximum mixing between the $\boldsymbol{\mu}$ and electron neutrinos, one must have

$$\boldsymbol{\Delta E = \frac{m^2_{\nu,\mu} - m^2_{\nu,e}}{2|p|}}. \tag{9.26}$$

Hence, when you read articles on solar neutrinos, the results are not plotted as functions of the neutrino masses, but as functions as the difference in neutrino masses squared.

9.3 Dressing Particles and Vertices

A photon can travel between galaxies, but during that time it isn't purely a photon. Sometimes it morphs into an electron positron pair, a $\boldsymbol{t\bar{t}}$ pair, or some very complicated mixture of particles and anti-particles. Similarly, if one creates an electron, by the time it interacts it will have morphed into a coherent mixture of innumerable states, the most common being the electron plus a photon. If one considers a creation operator $\boldsymbol{a^\dagger(k)}$, one needs to ask the question whether the operator is creating an eigenstate of the entire Hamiltonian, that includes all the mixtures, or is merely the particle without any of the accompanying cloud. We consider the particle as being "dressed", when the cloud is included, or "bare" when it is in the Fock state of just the particle by itself. If the particle is in an eigenstate, then one can consider it being dressed, but if it exists only transiently, it may be less dressed, or at least under-dressed. Since no process involves a transient time with zero duration, all particles are dressed at some level.
The particles being dressed affects their properties. In the vacuum this includes their masses and their couplings, i.e., their charges. In the context of perturbation theory, the dressing is described by the self-energy. Every diagram that begins and ends with the bare particle, but never has an intermediate state involving the bare particle, contributes to the self-energy. For bosons and fermions, these have the form,

$$\begin{aligned} \boldsymbol{-iG(k)} &= \boldsymbol{\frac{i}{p^2 - m_0^2 - \Pi(k) + i\epsilon}}, \\ \boldsymbol{-iS_F(p)} &= \boldsymbol{\frac{i}{\not{p} - m_0 - \Sigma(p) + i\epsilon}}. \end{aligned} \tag{9.27}$$

The self-energies are $\boldsymbol{\Pi(k)}$ for bosons, which has dimensions of energy squared, and $\boldsymbol{\Sigma(p)}$ for fermions, which has dimensions of energy. For vector particles, $\boldsymbol{\Pi}$ changes to $\boldsymbol{\Pi^{\mu\nu}}$.

For on-shell particles, the self-energy does two things. First, it changes the location of the pole, from $p^2 = m_0^2$ to $p^2 = m^2$, where m is the mass of the dressed particle. Since one never goes to the infinitesimally short-lived limit, the bare mass serves only as a straw man. Secondly, given that Π and Σ depend on k_0 and p_0, the self-energy affects the strength of the pole. If the strength of the pole was previously unity, adding the self-energy changes it to some new factor Z. For charged scalar bosons, the strength of the pole is:

$$Z = (d/dp^2)\left[p^2 - m_0^2 - \Pi(p^2)\right]. \tag{9.28}$$

Thus, after correcting for the self-energy the propagator would have the form,

$$G(p) = \frac{-1}{Z(p^2 - m^2) + \mathcal{O}(p^2 - m^2)^2 \cdots}. \tag{9.29}$$

One could have imagined that the original Lagrangian was defined with an extra piece

$$\Delta\mathcal{L} = (1 - Z)\left[-\partial_\mu \phi^* \partial^\mu \phi \cdots\right], \tag{9.30}$$

which would have resulted in a denominator without the factor Z above. This would be how one would effectively renormalize the fields so that the field operators effectively refer to dressed particles rather than bare ones, and Z is often referred to as a renormalization factor. Crudely, $1/Z$ describes the fraction of the dressed state that is the bare particle.
The most disquieting aspect of the procedure concerns dressings with infinite contributions, i.e., ultra-violet divergences. This derives from the integrals over loop momenta, where the internal momenta are integrated over an infinite range. The sickest diagram comes from dressing a boson with a fermion-antifermion loop.
The integral for the self energy,

$$\Pi^{\mu\nu} = e^2 \int \frac{d^4q}{(2\pi)^4} \frac{\mathrm{Tr}\gamma^\mu(\not{q} + m)\gamma^\nu(\not{q} - \not{k} + m)}{[q^2 - m^2 + i\epsilon][(q-k)^2 - m^2 + i\epsilon]}, \tag{9.31}$$

is quadratically divergent for large q. This leads to an infinite renormalization factor. There are two common procedures for sweeping such divergences under the rug, Pauli-Villars regularization and dimensional regularization. To better understand the methods requires significant investment and is beyond the scope of this course. Nonetheless, one can understand the issues it addresses without necessarily launching into the methods. Once the integrals have been regularized (for example in Pauli-Villars one simply defines a cutoff in momentum to make the contribution finite, then adds counter terms to the Lagrangian to cancel the contributions) the propagator behaves like a simple bare propagator for on-shell particles. However, once the particles are off-shell, the momentum dependence of the self energy comes back into play. For the photon, this effectively leads to a coupling that depends on the off-shellness (k^2 in the diagram above).
Dressing vertices is intrinsically more complicated since the diagrams involve at least three lines. Consider the lowest order correction to the electromagnetic coupling of an electron to a photon,

One might first expect that dressing the vertex would lead to a different effective charge. However, this would be most distressful. Since the correction to the vertex could depend on the mass of the incoming particle, one might expect that various particles would have different effective

couplings. This would mean that an electron might interact with the electromagnetic field with a different charge than a muon. Clearly this is experimentally untrue.
The solution to the worry above comes from the Ward Identity. The Ward identity relates the correction to the vertex to the correction to the self energy to ALL orders. For instance, consider a charged scalar field. The vertex to lowest order is simply is $2ek^\mu A_\mu$. Now consider the correction to the self-energy. Any contributing diagram has a line from the incoming charged particle that can be traced through the diagram from beginning to end that carries the incoming charge. If the incoming particle had momentum p, the internal lines that were part of the trace would all have momenta $p' = p + \sum_i Q_i$, where the momentum transfers to the traced line would be Q_i at each of the vertices encountered up to that point. By the end of the diagram, the last exchange would return the momentum back to its original value p. Further, p does not appear in any of the other internal lines.
The Ward Identity works by first noticing that if you begin with the collection of self-energy diagrams, you can generate the collection of all vertices with $k \to 0$ by replacing each line carrying the charge with

$$G_0(p') \to eG_0(p')(-2p'^\mu)G_0(p'). \tag{9.32}$$

Now take the derivative of Π with respect to p. This involves replacing the segments, one at a time, of the traced line with the derivative of the propagator,

$$G_0(p') \to \frac{d}{dp}G_0(p') = \frac{d}{dp'}G_0(p'). \tag{9.33}$$

The Ward identity comes from noticing that

$$\frac{d}{dp'_\mu}G_0(p') = G_0(p')(-2p'^\mu)G_0(p'). \tag{9.34}$$

Thus, by taking the derivative of $\Pi(p)$ with respect to p one seems to pick up all the possible vertices, except those that connect to internal loops. However, those loops all vanish in the limit of small exchanged momentum. This comes from each loop involving both a positive and negative particle which give opposite contributions.
After some consideration, one would think that the effective vertex would be of the form,

$$\Gamma^\mu(p, k \to 0, p+k) = e2p^\mu \frac{d}{dp^2}\left[p^2 - \Pi(p^2)\right]. \tag{9.35}$$

Since Π depends on the type of particle coming in, one would then think the charge, Γ^0, would depend on what type of particle one had. However, the effective charge can be written as

$$\Gamma^\mu(p, k \to 0, p) = 2Zp^\mu, \tag{9.36}$$

where Z is the strength of the propagator's pole as seen in Eq. (9.28). For renormalized propagators, one can replace the propagator to one without the factor of $1/Z$, but then add a factor of $1/Z^{1/2}$ to each line coming into a vertex. This exactly cancels the factor of Z in the vertex, making the effective vertex once again $2ep^\mu$, independent of the particles mass. Thus, all particles have the same effective charge, regardless of their mass, or how they interact with other particles.

The photon is also dressed according to its self energy. However, the photon couples to the charged particle only when the photon is in the Fock state of a single photon. This is accounted for by the renormalization constant $Z_\gamma(k)$, which is the strength of pole in the photon propagator. This correction applies to all particles equally. The loops involved in the photon propagator are referred to as polarization corrections to the vacuum, as they involve positive and negative pairs which you might think screen the vacuum. However, more physically, the pairs reduce the chance the photon is in a state that is capable of connecting to a charged particle. The effective coupling constant is $e^2 = e_0^2/Z_\gamma(k \to 0)$. The $k-$ dependence of $Z_\gamma(k)$ is responsible for the running coupling constant in QCD. Even in QED, the effective coupling constant change from $e^2/\hbar c = 1/137.036$ for $k = 0$ to $\approx 1/129$ for off-shellness near the W and Z masses.
The arguments above are only true for photons with vanishing momentum transfer. Once the momentum transfer is finite, the vertex corrections can no longer be ignored. This includes the coupling of the magnetic field to the spin. Since the magnetic field comes from taking derivatives w.r.t. A, this coupling is proportional to both k and A. The vertex corrections then change the g factor from being precisely equal to two, to being slightly different. Thus, the factor $g = 2 + \mathcal{O}\alpha$. The precise measurement of $g - 2$ has thus become the most demanding test of electromagnetic field theory. There is a race between being able to perform ever more accurate measurements, with theories that include ever more diagrams.

9.4 Homework Problems

1. Prove the Ward identity for fermions by showing that

$$\frac{d}{dp_\mu} S_F(p) = S_F(p)\gamma^\mu S_F(p),$$

where $S_F(p) = -(\not{p} + m)/(p^2 - m^2 + i\epsilon)$.

9.5 Vertices for electroweak interactions

Inclusion of leptons and quarks

Propagator: $\xrightarrow{\quad p \quad}$ $\dfrac{\mathrm{i}}{\not{p} - m_i + \mathrm{i}\varepsilon}$.

Vertices for leptons: $l = (\mathrm{e}, \mu, \tau)$, $\nu_l = (\nu_\mathrm{e}, \nu_\mu, \nu_\tau)$

for quarks q: p = (u, c, t), n = (d, s, b) with the CKM mixing matrix U_{pn} of eqn (12.39).