Relativistic Quantum Mechanics

Andreas Brandhuber

Department of Physics
Queen Mary, University of London
Mile End Road, London, E1 4NS
United Kingdom
email: a.brandhuber@qmul.ac.uk

Abstract

Lecture notes for the course Relativistic Quantum Mechanics MSci 4241/QMUL PHY-414. (These are preliminary notes and may contain typos, so use with care!)

1 Symmetries and Angular Momentum

1.1 Symmetries and Conservation Laws

In Quantum Mechanics (QM), for an observable A without explicit time dependence[1] we have

$$i\hbar\frac{d\langle A\rangle}{dt} = \langle\Psi|[\widehat{A},\widehat{H}]|\Psi\rangle\,, \tag{1.1}$$

for any state $|\Psi\rangle$ and where

$$\langle A\rangle = \langle\Psi|\widehat{A}|\Psi\rangle\,, \tag{1.2}$$

is the expectation value of A. If

$$[\widehat{A},\widehat{H}] = 0\,, \tag{1.3}$$

then

$$\frac{d\langle A\rangle}{dt} = 0\,, \tag{1.4}$$

and we say that A is a conserved quantitity or constant of motion.

A symmetry is a transformation on the coordinates of a system which leaves the Hamiltonian H invariant. We shall see that conservation laws are the consequence of symmetries of a system. Symmetries are very powerful since they can be used to derive results for a system even when we do not know the details of the dynamics involved.

1.2 Translational Invariance

First consider a single particle. If $\vec{x}$ is the position vector, then a translation is the operation

$$\vec{x} \to \vec{x}' = \vec{x} + \vec{a}\,. \tag{1.5}$$

If $\widehat{H}$ is invariant then

$$\widehat{H}(\vec{x}') = \widehat{H}(\vec{x}+\vec{a}) = \widehat{H}(\vec{x})\,. \tag{1.6}$$

For an infinitesimal displacement we can make a Taylor expansion[2]

$$\widehat{H}(\vec{x}+\vec{a}) \cong \widehat{H}(\vec{x}) + \vec{a}\cdot\vec{\nabla}\widehat{H}(\vec{x})\,, \tag{1.7}$$

ignoring higher powers of $\vec{a}$. Thus, if $\widehat{H}$ is invariant,

$$0 = \widehat{H}(\vec{x}+\vec{a}) - \widehat{H}(\vec{x}) = \vec{a}\cdot\vec{\nabla}\widehat{H}(\vec{x})\,. \tag{1.8}$$

[1] i.e. the QM operator $\widehat{A}$ corresponding to the observable A obeys $\partial\widehat{A}/\partial t = 0$.

[2] The symbol $\cong$ indicates that we only expand to first order in $\vec{a}$ and suppress all higher order terms.

In general for the momentum operator $\widehat{\vec{P}}$ and any other operator $\widehat{O}(\vec{x})$ we have

$$\begin{aligned}[\widehat{\vec{P}},\widehat{O}]\Psi &= [-i\hbar\vec{\nabla},\widehat{O}]\Psi \\ &= -i\hbar\vec{\nabla}(\widehat{O}\Psi) - \widehat{O}(-i\hbar\Psi) \\ &= -i\hbar(\vec{\nabla}\widehat{O})\Psi\,, \end{aligned} \tag{1.9}$$

where we have suppressed the explicit $\vec{x}$ and t dependence. Since eqn. (1.9) is true for arbitrary wavefunctions Ψ

$$[\widehat{\vec{P}},\widehat{O}(\vec{x})] = -i\hbar\vec{\nabla}\widehat{O}(\vec{x})\,. \tag{1.10}$$

In particular for $\widehat{O} = \widehat{H}$

$$[\widehat{\vec{P}},\widehat{H}] = -i\hbar\vec{\nabla}\widehat{H}\,. \tag{1.11}$$

Now,

$$0 = -i\hbar\vec{a}\cdot\vec{\nabla}\widehat{H} = \vec{a}\cdot[\widehat{\vec{P}},\widehat{H}] \tag{1.12}$$

and since this is true for an arbitrary displacement vector $\vec{a}$ we find

$$[\widehat{\vec{P}},\widehat{H}] = 0\,. \tag{1.13}$$

We conclude that momentum is a conserved quantity $\frac{d\langle\vec{P}\rangle}{dt} = 0$ if the $\widehat{H}$ is translationally invariant.

Take e.g.

$$\widehat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x}) \tag{1.14}$$

For the translation, $\vec{x}' = \vec{x} + \vec{a}$, in particular $x' = x + a$

$$\Rightarrow \frac{\partial}{\partial x} = \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} = \frac{\partial}{\partial x'} \tag{1.15}$$

and similarly for y and z. Thus $\vec{\nabla}$ and $\vec{\nabla}^2$ are invariant under translations and

$$\begin{aligned}\widehat{H}(\vec{x}') &= \widehat{H}(\vec{x}+\vec{a}) = -\frac{\hbar^2}{2m}\vec{\nabla}'^2 + V(\vec{x}') \\ &= -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x}+\vec{a})\,. \end{aligned} \tag{1.16}$$

Hence, for a translationally invariant Hamiltonian we must require

$$V(\vec{x}+\vec{a}) = V(\vec{x})\,, \tag{1.17}$$

which is only true for a (trivial) constant potential, i.e. for a free particle. Thus, the momentum of a free particle is conserved in QM in the sense

$$\frac{d\langle\vec{P}\rangle}{dt} = 0\,. \tag{1.18}$$

Consider now a two particle system (easily generalised to N particles). If the two particles have position vectors $\vec{x}_1$ and $\vec{x}_2$, the invariance condition for the translation of the system through $\vec{a}$ reads

$$\widehat{H}(\vec{x}_1, \vec{x}_2) = \widehat{H}(\vec{x}_1 + \vec{a}, \vec{x}_2 + \vec{a})\,. \tag{1.19}$$

Then for an infinitesimal translation

$$\widehat{H}(\vec{x}_1 + \vec{a}, \vec{x}_2 + \vec{a}) \cong \widehat{H}(\vec{x}_1, \vec{x}_2) + \vec{a} \cdot \vec{\nabla}_1 \widehat{H}(\vec{x}_1, \vec{x}_2) + \vec{a} \cdot \vec{\nabla}_2 \widehat{H}(\vec{x}_1, \vec{x}_2)\,, \tag{1.20}$$

we find that translational invariance implies

$$0 = \vec{a} \cdot (\vec{\nabla}_1 + \vec{\nabla}_2)\widehat{H}(\vec{x}_1, \vec{x}_2)\,, \tag{1.21}$$

and the total momentum operator is

$$\widehat{\vec{P}} = \widehat{\vec{P}}_1 + \widehat{\vec{P}}_2\,, \tag{1.22}$$

where

$$\widehat{\vec{P}}_1 = -i\hbar\vec{\nabla}_1 \text{ and } \widehat{\vec{P}}_2 = -i\hbar\vec{\nabla}_2\,. \tag{1.23}$$

Identical to the one particle case, for any operator $\widehat{O}(\vec{x}_1, \vec{x}_2)$,

$$[\widehat{\vec{P}}_1, \widehat{O}(\vec{x}_1, \vec{x}_2)] = -i\hbar\vec{\nabla}_1\widehat{O}(\vec{x}_1, \vec{x}_2) \text{ and } [\widehat{\vec{P}}_2, \widehat{O}(\vec{x}_1, \vec{x}_2)] = -i\hbar\vec{\nabla}_2\widehat{O}(\vec{x}_1, \vec{x}_2)\,, \tag{1.24}$$

and so

$$[\widehat{\vec{P}}, \widehat{O}(\vec{x}_1, \vec{x}_2)] = -i\hbar(\vec{\nabla}_1 + \vec{\nabla}_2)\widehat{O}(\vec{x}_1, \vec{x}_2)\,. \tag{1.25}$$

This is true in particular for $\widehat{O} = \widehat{H}$. Thus,

$$0 = -i\hbar\vec{a} \cdot (\vec{\nabla}_1 + \vec{\nabla}_2)\widehat{H} = \vec{a} \cdot [\widehat{\vec{P}}, \widehat{H}]\,. \tag{1.26}$$

Since this must be true for arbitrary translation vector $\vec{a}$, we have

$$[\widehat{\vec{P}}, \widehat{H}] = 0\,, \tag{1.27}$$

and total momentum is conserved in the sense of QM i.e. $\frac{d\langle\vec{P}\rangle}{dt} = 0$.

1.3 Rotational Invariance

Just as translational invariande is associated with conservation of momentum, it turns out that rotational invariance is associated with conservation of angular momentum (AM).

Take spherical polar coordinates and take the axis of rotation to be the z-axis. Specify the position vector $\vec{x}$ in spherical polar coordinates (r, θ, ϕ) of the point. Then the

symmetry operation $\vec{x} \to \vec{x}'$ corresponding to a rotation by an angle α about the $z-$axis is

$$(r, \theta, \phi) \to (r', \theta', \phi') = (r, \theta, \phi + \alpha)\,. \tag{1.28}$$

For the Hamiltonian to be invariant under rotations about the z-axis

$$\begin{aligned} & \widehat{H}(\vec{x}') = \widehat{H}(\vec{x}) \\ \Longleftrightarrow \quad & \widehat{H}(r', \theta', \phi') = \widehat{H}(r, \theta, \phi) \\ \Longleftrightarrow \quad & \widehat{H}(r, \theta, \phi + \alpha) = \widehat{H}(r, \theta, \phi)\,. \end{aligned} \tag{1.29}$$

For an infinitesimal rotation

$$\widehat{H}(r, \theta, \phi + \alpha) \cong \widehat{H}(r, \theta, \phi) + \alpha \frac{\partial}{\partial \phi} \widehat{H}(r, \theta, \phi)\,, \tag{1.30}$$

and, hence, for invariance of $\widehat{H}$

$$\begin{aligned} 0 \quad &= \quad \widehat{H}(r, \theta, \phi + \alpha) - \widehat{H}(r, \theta, \phi) = \alpha \frac{\partial}{\partial \phi} \widehat{H}(r, \theta, \phi) \\ &\longrightarrow \quad \frac{\partial}{\partial \phi} \widehat{H}(r, \theta, \phi) = 0\,. \end{aligned} \tag{1.31}$$

The $z-$component of the orbital AM operator written in spherical coordinates is[3]

$$\widehat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \tag{1.32}$$

In general, for any other operator $\widehat{O}$,

$$[\widehat{L}_z, \widehat{O}]\Psi = -i\hbar [\frac{\partial}{\partial \phi}, \widehat{O}] = -i\hbar (\frac{\partial}{\partial \phi}(\widehat{O}\Psi) - \widehat{O}\frac{\partial}{\partial \phi}\Psi) = -i\hbar \frac{\partial \widehat{O}}{\partial \phi}\Psi\,. \tag{1.33}$$

This is true for arbitrary wavefuntions Ψ, thus

$$[\widehat{L}_z, \widehat{O}] = -i\hbar \frac{\partial \widehat{O}}{\partial \phi}\,, \tag{1.34}$$

and in particular for $\widehat{O} = \widehat{H}$,

$$[\widehat{L}_z, \widehat{H}] = -i\hbar \frac{\partial \widehat{H}}{\partial \phi}\,. \tag{1.35}$$

If the Hamiltonian is invariant under rotation about the z-axis, we now conclude that

$$[\widehat{L}_z, \widehat{H}] = 0\,. \tag{1.36}$$

[3] In cartesian coordinates $\vec{x} = (x, y, z)$: $\widehat{L}_z = (\widehat{\vec{x}} \times \widehat{\vec{P}})_z = -i\hbar (x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x})$. The other components, $\widehat{L}_x$ and $\widehat{L}_y$, can be obtained by cyclic permutation of (x, y, z).

We can define angles ϕ_x and ϕ_y analogous to $\phi_z \equiv \phi$ for rotations about the x and y-axis. If $\widehat{H}$ is also invariant under rotations about the x and y-axis we will conclude that

$$[\widehat{L}_x, \widehat{H}] = [\widehat{L}_y, \widehat{H}] = [\widehat{L}_z, \widehat{H}] = 0 \text{ i.e. } [\widehat{\vec{L}}, \widehat{H}] = 0\,. \tag{1.37}$$

Thus AM is a constant of motion

$$\frac{d\langle \vec{L} \rangle}{dt} = 0\,. \tag{1.38}$$

Any rotation can be built out of successive rotations about the x, y and z-axis.

Whenever $\widehat{H}$ is invariant under arbitrary rotations the AM $\vec{L}$ is a conserved quantity. We must construct such an $\widehat{H}$ out of scalars i.e. invariants under rotations. The simplest examples of scalars are the magnitude (length) of a vector or the scalar product of two vectors. For example, consider the Hamiltonian for a particle moving in a central potential such as the Coulomb potential

$$\widehat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(|\vec{x}|)\,. \tag{1.39}$$

Because $\vec{\nabla}^2 = \vec{\nabla}\cdot\vec{\nabla}$ and $|\vec{x}|$ are scalars, so is the Hamiltonian and orbital AM is conserved.

This discussion generalizes immediately to two (or more) particles:

For 2 particles, the Hamiltonian is a function of two sets of spherical coordinates (r_1, θ_1, ϕ_1) and (r_2, θ_2, ϕ_2) so that

$$\widehat{H} = \widehat{H}(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2)\,. \tag{1.40}$$

The invariance condition for rotation of the system by an angle α about the z-axis is

$$\widehat{H}(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) = \widehat{H}(r_1, \theta_1, \phi_1 + \alpha; r_2, \theta_2, \phi_2 + \alpha)\,, \tag{1.41}$$

and for an infinitesimal rotation

$$\widehat{H}(r_1, \theta_1, \phi_1 + \alpha; r_2, \theta_2, \phi_2 + \alpha) \cong \widehat{H}(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) + \alpha\partial_{\phi_1}\widehat{H} + \alpha\partial_{\phi_2}\widehat{H}\,, \tag{1.42}$$

where we defined $\partial_\phi \equiv \frac{\partial}{\partial\phi}$. Invariance of $\widehat{H}$ gives

$$\begin{aligned} 0 &= \widehat{H}(r_1, \theta_1, \phi_1 + \alpha; r_2, \theta_2, \phi_2 + \alpha) - \widehat{H}(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) \\ &= \alpha(\partial_{\phi_1} + \partial_{\phi_2})\widehat{H} \\ \longrightarrow\ & (\partial_{\phi_1} + \partial_{\phi_2})\widehat{H} = 0\,. \end{aligned} \tag{1.43}$$

The z-components of the orbital AM operator for the two particles are

$$\widehat{L}_{1z} = -i\hbar\partial_{\phi_1}\ ,\ \widehat{L}_{2z} = -i\hbar\partial_{\phi_2}\,. \tag{1.44}$$

The z-component of the total orbital AM is:

$$\widehat{L}_z = \widehat{L}_{1z} + \widehat{L}_{2z}\,. \tag{1.45}$$

Much as before, for any operator O,

$$[\widehat{L}_z, \widehat{O}] = -i\hbar(\partial_{\phi_1} + \partial_{\phi_2})\widehat{O}\,, \tag{1.46}$$

and in particular

$$[\widehat{L}_z, \widehat{H}] = -i\hbar(\partial_{\phi_1} + \partial_{\phi_2})\widehat{H}\,. \tag{1.47}$$

If the Hamiltonian is invariant under rotations about the z-axis, we now conclude that

$$[\widehat{L}_z, \widehat{H}] = 0\,. \tag{1.48}$$

By also considering rotations about the x and y-axis we conclude that if $\widehat{H}$ is invariant under arbitrary rotations then

$$[\widehat{\vec{L}}, \widehat{H}] = 0\,, \tag{1.49}$$

so total AM is a constant of motion

$$\frac{d\langle\vec{L}\rangle}{dt} = 0\,. \tag{1.50}$$

As usual the orbital AM operators obey the ($SO(3)$ or $SU(2)$) algebra

$$[\widehat{L}_x, \widehat{L}_y] = i\hbar\widehat{L}_z \;\;,\;\; [\widehat{L}_x, \widehat{L}_y] = i\hbar\widehat{L}_z \;\;,\;\; [\widehat{L}_x, \widehat{L}_y] = i\hbar\widehat{L}_z\,. \tag{1.51}$$

Consequently, it is only possible to know simultaneously the values of one component $\vec{L}$ and $\vec{L}^2$, for example, L_z and $\vec{L}^2$. Therefore, we may take the conserved quantities to be L_z and $\vec{L}^2$.

An atom with atomic number Z with Coulomb forces between the nucleus and the electrons and between the electrons is an example of a system with the necessary rotational invariance for conservation of AM. In this case:

$$\widehat{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{Z}\vec{\nabla}_i^2 - \sum_{i=1}^{Z}\frac{Ze^2}{4\pi\epsilon_0|\vec{x}_i|} + \sum_{i,j=1,i<j}^{Z}\frac{e^2}{4\pi\epsilon_0|\vec{x}_i - \vec{x}_j|}\,, \tag{1.52}$$

where $\vec{\nabla}_i$ acts on the coordinates of the i-th electron, and $\vec{x}_i$ is the position vector of the i-th electron relative to the nucleus. As before, $\vec{\nabla}_i^2$ and $|\vec{x}_i|$ are scalars (rotationally invariant) and so is $|\vec{x}_i - \vec{x}_j|$, so that $\widehat{H}$ is rotationally invariant.

Remark: Generally for a system with spin it is the total AM

$$\widehat{\vec{J}} = \widehat{\vec{L}} + \widehat{\vec{S}} \tag{1.53}$$

that commutes with the Hamiltonian

$$[\widehat{\vec{J}}, \widehat{H}] = 0 \tag{1.54}$$

if $\widehat{H}$ is rotationally invariant, and then $\langle J_z \rangle$ and $\langle \vec{J}^2 \rangle$ are constants of motion.

For the atomic (non-relativistic) Hamiltonian above, $\langle L_z \rangle$ and $\langle \vec{L}^2 \rangle$ are also constants of motion, because the spin does not appear explicitly in the Hamiltion. This is a manifestation of the fact that spin is an effect of Special Relativity as we will see later in the course. However, if we include the Spin-Orbit interaction due to Relativistic effects

$$\widehat{H}_{Spin-Orbit} = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV}{dr} \widehat{\vec{L}} \cdot \widehat{\vec{S}}\,, \tag{1.55}$$

for an electron moving in a central potential $V(r)$, then $\widehat{\vec{L}}$ does not commute with the complete Hamiltonian. Then, only $\langle J_z \rangle$ and $\langle \vec{J}^2 \rangle$ are conserved quantities.

1.4 Addition of Angular Momentum

For a particle with orbital AM operator $\widehat{\vec{L}}$ and spin angular momentum operator $\widehat{\vec{S}}$ we often need to construct the eigenstates of the total AM operator

$$\widehat{\vec{J}} = \widehat{\vec{L}} + \widehat{\vec{S}}\,, \tag{1.56}$$

in terms of eigenstates of $\widehat{\vec{L}}$ and $\widehat{\vec{S}}$. (For convenience we will drop the $\widehat{\ }$ for operators from now on.) More precisely, we need to construct the simultaneous eigenstates $|j, m\rangle$ of $\vec{J}^2$ and J_z in terms of simultaneous eigenstates[4] $|l, m_l\rangle$ of $\vec{L}^2$ and L_z and $|s, m_s\rangle$ of $\vec{S}^2$ and S_z.

Recall from QM that for eigenvalue of $\vec{L}^2$ equal to $l(l+1)\hbar^2$ with l an integer, the eigenvalues of L_z are $m_l\hbar$ with m_l taking values $m_l = -l, -l+1, \ldots, l-1, l$ and similarly for $\vec{S}$ and $\vec{J}$ where s and j are integer or half-integer.

Also for a 2 particle system, we often need simultaneous eigenstates $|jm\rangle$ of the total AM operators $\vec{J}^2$ and J_z in terms of simultaneous eigenstates $|j_1 m_1\rangle$ of $\vec{J}_1^2$ and J_{1z} and $|j_2 m_2\rangle$ of $\vec{J}_2^2$ and J_{2z}.

In the following we assume that all eigenstates are "orthonormalized" in the familiar fashion, e.g. $\langle jm|j'm'\rangle = \delta_{jj'}\delta_{mm'}$.

[4] In QM we encounter the eigenstates $|l, m_l\rangle$ when the Schrödinger equation is solved in spherical coordinates. Usually they are denoted as spherical harmonics $Y_{l,m_l}(\theta, \phi)$ but their explicit form will not be needed in the following.

Clebsch-Gordon coefficients are defined to be the coefficients in the expansion

$$|jm\rangle = \sum_{m_1,m_2} C(j_1 m_1, j_2 m_2|jm)|j_1 m_1\rangle|j_2 m_2\rangle\,. \tag{1.57}$$

Range of Values of j and m

Consider $\vec{J} = \vec{J_1} + \vec{J_2}$ ($\vec{J} = \vec{L} + \vec{S}$ is exactly similar):

- We have
$$\vec{J} = \vec{J_1} + \vec{J_2} \longrightarrow J_z = J_{1z} + J_{2z}\,, \tag{1.58}$$
thus for any simultaneous eigenstate $|\Psi\rangle$ of J_{1z} and J_{2z} we have
$$\begin{aligned} J_z|\Psi\rangle &= (J_{1z} + J_{2z})|\Psi\rangle \\ &\longrightarrow m\hbar|\Psi\rangle = (m_1\hbar + m_2\hbar)|\Psi\rangle \\ &\longrightarrow m = m_1 + m_2\,. \end{aligned} \tag{1.59}$$
Now $m = m_1 + m_2$ is the eigenvalue of J_z which therefore runs from $-j$ to $+j$.

- No matter what j we construct from j_1 and j_2, we must have
$$m_u \leq m_{1u} + m_{2u}\,, \tag{1.60}$$
where $m_u\hbar$, $m_{1u}\hbar$ and $m_{2u}\hbar$ are the maximal eigenvalues of J_z, J_{1z} and J_{2z}, respectively. Since these maximal eigenvalues are just $j\hbar$, $j_1\hbar$ and $j_2\hbar$, we must have
$$j \leq j_1 + j_2\,. \tag{1.61}$$

- Suppose $j_1 > j_2$, then (without proof) $m_u \geq m_{1u} - m_{2u}$,
$$\Rightarrow j \geq j_1 - j_2\,. \tag{1.62}$$
Similarly, if $j_2 > j_1$,
$$\Rightarrow j \geq j_2 - j_1\,. \tag{1.63}$$
In general:
$$j \geq |j_1 - j_2| \tag{1.64}$$

Thus, the possible values of j that can be constructed from j_1 and j_2 are

$$j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, j_1 + j_2 - 1, j_1 + j_2\,. \tag{1.65}$$

This is also known as the triangle condition.

Explicit Construction of Clebsch-Gordon Coefficients

$\underline{j_1 = \frac{1}{2},\, j_2 = 1}$

We shall need to know $J_-|j,m\rangle$ where $J_\pm = J_x \pm iJ_y$ are the familiar raising and lowering operators. Note that they are the adjoint operators of each other, i.e. $J_+^\dagger = J_-$ and $J_-^\dagger = J_+$ and obey $J_+J_- = \vec{J}^2 - J_z^2 + \hbar J_z$.

$J_-|j,m\rangle$ has eigenvalue $J_z = (m-1)\hbar$, thus $J_-|j,m\rangle \propto |j,m-1\rangle$.

The constant of proportionality can be determined (without proof) with the result

$$J_-|j,m\rangle = \sqrt{(j-m+1)(j+m)}\hbar|j,m-1\rangle\,. \tag{1.66}$$

For the case $j_1 = \frac{1}{2}$, $j_2 = 1$, the possible values of j are $\frac{1}{2}$ and $\frac{3}{2}$. (Recall $j = |j_1 - j_2|, \ldots, j_1 + j_2 - 1, j_1 + j_2$.)

Begin by constructing the state with maximum value of j and maximum value of m, i.e. $|j = \frac{3}{2}, m = \frac{3}{2}\rangle$.

Since $m = m_1 + m_2$ and $m_1 = \pm\frac{1}{2}$ and $m_2 = -1, 0, +1$, there is only one way to get $m = \frac{3}{2}$, namely $\frac{3}{2} = \frac{1}{2} + 1$. Thus,

$$|j = \frac{3}{2}, m = \frac{3}{2}\rangle = |j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\rangle|j_2 = 1, m_2 = 1\rangle\,. \tag{1.67}$$

We can now use the lowering operator J_- to construct states with the same value for j but smaller values of m.

$$\begin{aligned} \vec{J} &= \vec{J}_1 + \vec{J}_2 \\ &\Rightarrow \quad J_- = (J_1)_- + (J_2)_-\,. \end{aligned} \tag{1.68}$$

From now on we will take natural units with $\hbar = 1$. Apply eqn. (1.68) to eqn. (1.67). The left hand side gives

$$J_-|j = \frac{3}{2}, m = \frac{3}{2}\rangle = \sqrt{(3/2 - 3/2 + 1)(3/2 + 3/2)}|j = \frac{3}{2}, m = \frac{1}{2}\rangle = \sqrt{3}|j = \frac{3}{2}, m = \frac{1}{2}\rangle\,, \tag{1.69}$$

whereas the right hand side gives

$$\begin{aligned} J_-|j=\frac{3}{2},m=\frac{3}{2}\rangle &= ((J_1)_- + (J_2)_-)|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=1\rangle \\ &= ((J_1)_-|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle)|j_2=1,m_2=1\rangle + \\ &\quad |j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle((J_2)_-|j_2=1,m_2=1\rangle) \\ &= \sqrt{(1/2-1/2+1)(1/2+1/2)}|j_1=\frac{1}{2},m_1=-\frac{1}{2}\rangle|j_2=1,m_2=1\rangle + \\ &\quad \sqrt{(1-1+1)(1+1)}|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=0\rangle\,, \end{aligned} \tag{1.70}$$

thus,

$$\begin{aligned} J_-|j=\frac{3}{2},m=\frac{3}{2}\rangle &= ((J_1)_- + (J_2)_-)|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=1\rangle \\ &= |j_1=\frac{1}{2},m_1=-\frac{1}{2}\rangle|j_2=1,m_2=1\rangle + \\ &\quad \sqrt{2}|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=0\rangle\,. \end{aligned} \tag{1.71}$$

Comparing left and right hand side we obtain,

$$\begin{aligned} |j=\frac{3}{2},m=\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}}|j_1=\frac{1}{2},m_1=-\frac{1}{2}\rangle|j_2=1,m_2=1\rangle + \\ &\quad \frac{\sqrt{2}}{\sqrt{3}}|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=0\rangle\,. \end{aligned} \tag{1.72}$$

Now we can apply J_- again to this state, and after a similar calculation as before we obtain

$$\begin{aligned} |j=\frac{3}{2},m=-\frac{1}{2}\rangle &= \frac{\sqrt{2}}{\sqrt{3}}|j_1=\frac{1}{2},m_1=-\frac{1}{2}\rangle|j_2=1,m_2=0\rangle + \\ &\quad \sqrt{\frac{1}{3}}|j_1=\frac{1}{2},m_1=\frac{1}{2}\rangle|j_2=1,m_2=-1\rangle\,. \end{aligned} \tag{1.73}$$

Finally we can apply J_- a third (and last) time to obtain

$$|j=\frac{3}{2},m=-\frac{3}{2}\rangle = |j_1=\frac{1}{2},m_1=-\frac{1}{2}\rangle|j_2=1,m_2=-1\rangle \tag{1.74}$$

We have now constructed *all* $j=\frac{3}{2}$ states. Acting with J_- again gives zero and terminates the procedure, which is consistent with the fact that $m=-5/2$ would be outside of the allowed range.

How can we obtain the $j = \frac{1}{2}$ states?

Start from the state with $j = \frac{1}{2}$ and the maximum value of m (for that j) i.e. $|j = \frac{1}{2}, m = \frac{1}{2}\rangle$. This state has the same value of m as $|j = \frac{3}{2}, m = \frac{1}{2}\rangle$ and must be a linear combination of the same $|j_1, m_1\rangle|j_2, m_2\rangle$ states. However, because it has different value of j it must be orthogonal to $|j = \frac{3}{2}, m = \frac{1}{2}\rangle$. (Distinct eigenvalues of a Hermitian Operator have orthogonal eigenfunctions.)

Write,

$$\begin{aligned} |j = \frac{1}{2}, m = \frac{1}{2}\rangle &= c_1|j_1 = \frac{1}{2}, m_1 = -\frac{1}{2}\rangle|j_2 = 1, m_2 = 1\rangle + \\ & \quad c_2|j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\rangle|j_2 = 1, m_2 = 0\rangle\,. \end{aligned} \tag{1.75}$$

Orthogonality to $|j = \frac{3}{2}, m = \frac{1}{2}\rangle$ implies

$$\langle j = \frac{3}{2}, m = \frac{1}{2}|j = \frac{1}{2}, m = \frac{1}{2}\rangle = 0\,, \tag{1.76}$$

hence,

$$\begin{aligned} &\frac{c_1}{\sqrt{3}}\langle j_1 = \frac{1}{2}, m_1 = -\frac{1}{2}|j_1 = \frac{1}{2}, m_1 = -\frac{1}{2}\rangle\langle j_2 = 1, m_2 = 1|j_2 = 1, m_2 = 1\rangle + \\ &\frac{c_2\sqrt{2}}{\sqrt{3}}\langle j_1 = \frac{1}{2}, m_1 = \frac{1}{2}|j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\rangle\langle j_2 = 1, m_2 = 0|j_2 = 1, m_2 = 0\rangle + \\ &\frac{c_2}{\sqrt{3}}\langle j_1 = \frac{1}{2}, m_1 = -\frac{1}{2}|j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\rangle\langle j_2 = 1, m_2 = 1|j_2 = 1, m_2 = 0\rangle + \\ &\frac{c_1\sqrt{2}}{\sqrt{3}}\langle j_1 = \frac{1}{2}, m_1 = \frac{1}{2}|j_1 = \frac{1}{2}, m_1 = -\frac{1}{2}\rangle\langle j_2 = 1, m_2 = 0|j_2 = 1, m_2 = 1\rangle \\ &= 0\,. \end{aligned} \tag{1.77}$$

Due to "orthonormality" of the eigenstates, the third and fourth line of this equation vanish and from the rest we get

$$\frac{c_1}{\sqrt{3}} + \sqrt{\frac{2}{3}}c_2 = 0 \Rightarrow c_1 = -\sqrt{2}c_2\,. \tag{1.78}$$

Furthermore, the state must be normalized

$$\Rightarrow \langle j = \frac{1}{2}, m = \frac{1}{2}|j = \frac{1}{2}, m = \frac{1}{2}\rangle = 1 \tag{1.79}$$

hence,

$$\Rightarrow |c_1|^2 + |c_2|^2 = 1\,. \tag{1.80}$$

Take c_1 and c_2 real[5], then

$$\begin{aligned} 2c_2^2 + c_2^2 &= 3c_2^2 = 1 \\ &\Rightarrow \quad c_2 = \frac{1}{\sqrt{3}} \Rightarrow c_1 = -\sqrt{\frac{2}{3}}\,, \end{aligned} \tag{1.81}$$

so

$$\begin{aligned} |j=\frac{1}{2}, m=\frac{1}{2}\rangle &= -\sqrt{\frac{2}{3}}|j_1=\frac{1}{2}, m_1=-\frac{1}{2}\rangle|j_2=1, m_2=1\rangle \\ &\quad +\frac{1}{\sqrt{3}}|j_1=\frac{1}{2}, m_1=\frac{1}{2}\rangle|j_2=1, m_2=0\rangle\,. \end{aligned} \tag{1.82}$$

Using again the J_- operator we can also obtain the state $|j=\frac{1}{2}, m=-\frac{1}{2}\rangle$ with the result

$$\begin{aligned} |j=\frac{1}{2}, m=-\frac{1}{2}\rangle &= \sqrt{\frac{2}{3}}|j_1=\frac{1}{2}, m_1=\frac{1}{2}\rangle|j_2=1, m_2=-1\rangle \\ &\quad -\frac{1}{\sqrt{3}}|j_1=\frac{1}{2}, m_1=-\frac{1}{2}\rangle|j_2=1, m_2=0\rangle\,. \end{aligned} \tag{1.83}$$

1.5 The $O(4)$ Symmetry of the Hydrogen Atom

For the Hydrogen atom, ignoring small terms,

$$\widehat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(|\vec{x}|) \tag{1.84}$$

with $V(|\vec{x}|) = -\frac{e^2}{4\pi\epsilon_0|\vec{x}|}$. As shown in a previous Section, the Rotational Invariance of $\widehat{H}$ means that

$$[\widehat{\vec{L}}, \widehat{H}] = 0 \tag{1.85}$$

and hence orbital AM is a conserved quantity. (Remember $\widehat{L}_\pm = \widehat{L}_x \pm i\widehat{L}_y$.)

If we consider a state $|l, m_l\rangle$ with energy E, then

$$\begin{aligned} \widehat{H}\left(\widehat{L}_\pm\right)|l, m_l\rangle &= \widehat{L}_\pm\widehat{H}|l, m_l\rangle \\ &= \widehat{L}_\pm E|l, m_l\rangle \\ &= E\widehat{L}_\pm|l, m_l\rangle\,. \end{aligned} \tag{1.86}$$

Thus, $\widehat{L}_\pm|l, m_l\rangle \propto |l, m_l \pm 1\rangle$ also has energy E. Consequently, all $2l+1$ states with $m_l = -l, -l+1, \ldots, l-1, l$, in an AM multiplet with eigenvalue $l(l+1)\hbar^2$ of $\widehat{\vec{L}}^2$, are

[5]This fixes an irrelevant overall phase factor.

degenerate in energy. This degeneracy is a manifestation of the Rotational Symmetry of the Hydrogen atom.

It is known that for ${}^1_1H$, not only are all the states with the same AM Quantum Number l degenerate in energy, but also are all the states with $l = 0, 1, \ldots, n-1$ corresponding to the Principle Quantum Number n (the proof can be found in standard QM books). We want to show here that this is a consequence of a larger symmetry group that includes Rotational invariance.

In classical physics it can be shown that for a central potential $V(|\vec{x}|)$ not only is

$$\vec{L} = \vec{x} \times \vec{p}, \tag{1.87}$$

a constant of motion, but so also is the Lenz vector (in units where $m_e = 1$)

$$\vec{M} = \frac{1}{\sqrt{-8E}} \left(\vec{L} \times \vec{x} - \vec{x} \times \vec{L} + \frac{e^2}{2\pi\epsilon_0} \frac{\vec{x}}{|\vec{x}|} \,. \right) \tag{1.88}$$

Correspondingly, in QM both

$$\widehat{\vec{L}} = \widehat{\vec{x}} \times \widehat{\vec{p}}, \tag{1.89}$$

and

$$\widehat{\vec{M}} = \frac{1}{\sqrt{-8\widehat{H}}} \left(\widehat{\vec{L}} \times \widehat{\vec{x}} - \widehat{\vec{x}} \times \widehat{\vec{L}} + \frac{e^2}{2\pi\epsilon_0} \frac{\widehat{\vec{x}}}{|\vec{x}|} \,, \right) \tag{1.90}$$

commute with $\widehat{H}$. If we define the linear combinations

$$\widehat{\vec{I}} = \frac{1}{2}(\widehat{\vec{L}} + \widehat{\vec{M}}) \;\text{ and }\; \widehat{\vec{K}} = \frac{1}{2}(\widehat{\vec{L}} - \widehat{\vec{M}}) \tag{1.91}$$

it can be shown that $\widehat{\vec{I}}$ and $\widehat{\vec{K}}$ commute with each other and they behave like AM operators i.e.

$$[\widehat{I}_x, \widehat{I}_y] = i\widehat{I}_z \;\;,\;\; [\widehat{K}_x, \widehat{K}_y] = i\widehat{K}_z \,, \;\text{ plus cyclic permutations} \tag{1.92}$$

with $\hbar = 1$. This is referred to as the $O(4)$ or $SU(2) \times SU(2)$ algebra. With some effort it can be shown that the Hamiltonian can be written in the following simple form

$$\widehat{H} = -\frac{1}{4} \frac{1}{\widehat{\vec{I}}^2 + \widehat{\vec{K}}^2 + \frac{1}{2}} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \tag{1.93}$$

Because $\widehat{\vec{I}}$ and $\widehat{\vec{K}}$ each obey the AM algebra, the eigenvalues of $\widehat{\vec{I}}^2$ and $\widehat{\vec{K}}^2$ are of the form $i(i+1)$ and $k(k+1)$ with i and k integers (with $\hbar = 1$).

Thus the energy levels of the hydrogen atom are

$$E(i,k) = -\frac{1}{4} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left(i(i+1) + k(k+1) + \frac{1}{2} \right)^{-1} \,. \tag{1.94}$$

These energy levels have degeneracy $(2i+1)(2k+1)$.

Because of the properties of the triple scalar product $\vec{a}\cdot(\vec{b}\times\vec{c})$ we have

$$\widehat{\vec{L}}\cdot\widehat{\vec{M}}=0\,, \tag{1.95}$$

thus,

$$\widehat{\vec{I}}^{\,2}=\widehat{\vec{K}}^{\,2}=\frac{1}{4}\left(\widehat{\vec{L}}^{\,2}+\widehat{\vec{M}}^{\,2}\right) \tag{1.96}$$

and we must have $i=k$. If we write $n=2i+1=2k+1$:

$$\Rightarrow i(i+1)=\frac{n^2-1}{4} \tag{1.97}$$

Then the energy levels are

$$E_n=-\frac{1}{4}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\left(\frac{n^2-1}{4}+\frac{n^2-1}{4}+\frac{1}{2}\right)^{-1}=-\frac{1}{2n^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2. \tag{1.98}$$

As expected having set $m_e=1$ and $\hbar=1$ with Degeneracy n^2.

Moreover, because $\widehat{\vec{L}}=\widehat{\vec{I}}+\widehat{\vec{K}}$ and both $\widehat{\vec{I}}$ and $\widehat{\vec{K}}$ both behave like AM operators, using the rules for adding AM the possible values of l are $l=i+k, i+k-1,\ldots,|i-k|+1,|i-k|$. Setting $i=k=(n-1)/2$, we see that l runs from 0 to $n-1$, as required.

The Degeneracy is also the Expected Degeneracy because we have degenerate states with $l=0,1,\ldots,n-1$ and for each l, $m_l=-l,\ldots,l$. Thus there are

$$\sum_{l=0}^{n-1}(2l+1)=n^2 \tag{1.99}$$

degenerate states for the principle quantum number n.

2 Relativistic Quantum Mechanics

2.1 Four Vector Formalism

Any point of space-time is described by $x^\mu, \mu=0,1,2,3$ with $x^0=ct$, $x^1=x$, $x^2=y$ and $x^3=z$. Any mathematical object a^μ is called a four-vector if it transforms in the same fashion as x^μ under Lorentz transformations (LT's).

If a^μ and b^μ are four-vectors

$$a \cdot b \equiv a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 \tag{2.1}$$

is a scalar (invariant) under LT's. It has the same value for all observers just like

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = c^2 t^2 - x^2 - y^2 - z^2 \tag{2.2}$$

An important four-vector is the four-momentum

$$p^\mu \equiv \left(\frac{E}{c}, p_x, p_y, p_z\right) \equiv \left(\frac{E}{c}, \vec{p}\right) \tag{2.3}$$

where $\vec{p}$ is the 3-momentum. The corresponding Lorentz invariant (scalar) quantity is

$$\begin{aligned} p^2 \equiv p^\mu p_\mu &\equiv (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2 \\ &\Rightarrow \quad E^2 = \vec{p}^2 c^2 + m^2 c^4 \,. \end{aligned} \tag{2.4}$$

In QM the corresponding hermitian operators are

$$\begin{aligned} \widehat{p}^\mu &= \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, -i\hbar\frac{\partial}{\partial x}, -i\hbar\frac{\partial}{\partial y}, -i\hbar\frac{\partial}{\partial z}\right) \\ &\equiv \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, -i\hbar\vec{\nabla}\right) \\ \Rightarrow \widehat{p}^\mu &= \left(i\hbar\frac{\partial}{\partial x^0}, -i\hbar\vec{\nabla}\right) \,. \end{aligned} \tag{2.5}$$

If we define the differential operator 4-vectors, $\nabla^\mu = \left(\frac{\partial}{\partial x^0}, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right)$ and $\nabla_\mu = \left(\frac{\partial}{\partial x^0}, \vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right)$, then the 4-momentum operator is

$$\widehat{p}^\mu = i\hbar\nabla^\mu \,, \tag{2.6}$$

and the Lorentz invariant (scalar) constructed from ∇^μ is

$$\Box \equiv \nabla \cdot \nabla = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \,. \tag{2.7}$$

2.2 The Klein-Gordon Equation

In non-relativistic QM, the free Hamiltonian $H = E = \frac{\vec{p}^2}{2m}$ is quantised by the substitution

$$H \to i\hbar\frac{\partial}{\partial t} \;\;,\;\; \vec{p} \to -i\hbar\vec{\nabla} \tag{2.8}$$

to give the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\Psi\,. \tag{2.9}$$

A relativistic free particle has Hamiltonian

$$H = E = \sqrt{\vec{p}^2c^2 + m^2c^4} \tag{2.10}$$

and hence the same, naive substitution gives

$$i\hbar\frac{\partial\Psi}{\partial t} = \sqrt{m^2c^4 - \hbar^2c^2\vec{\nabla}^2}\Psi\,. \tag{2.11}$$

But what to do about the square root of the operator? One interpretation is to make a series expansion, but then we get a Hamiltonian with derivatives of arbitrarily high order.

A more sensible route is to start from $H^2 = \vec{p}^2c^2 + m^2c^4$ to get

$$\begin{aligned} -\hbar^2\frac{\partial^2}{\partial t^2}\Psi &= \left(-\hbar^2c^2\vec{\nabla}^2 + m^2c^4\right)\Psi \\ &\Rightarrow \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2\right)\Psi + \left(\frac{mc}{\hbar}\right)^2\Psi = 0 \\ &\Rightarrow \left(\Box + \left(\frac{mc}{\hbar}\right)^2\right)\Psi = 0 \end{aligned} \tag{2.12}$$

By analogy with the Schrödinger equation it is possible to derive a continuity equation for the Klein-Gordon (KG) equation

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{2.13}$$

where ρ is the probability density and $\vec{J}$ is the probability current.[6]

Derivation: Subtracting the following two equations

$$\begin{aligned} \Psi^*\left(\Box + \left(\frac{mc}{\hbar}\right)^2\right)\Psi = 0 \\ \Psi\left(\Box + \left(\frac{mc}{\hbar}\right)^2\right)\Psi^* = 0 \end{aligned} \tag{2.14}$$

gives

$$\Psi^*\nabla\cdot\nabla\Psi - \Psi\nabla\cdot\nabla\Psi^*\,. \tag{2.15}$$

[6]Integrating the probability density over a volume V bounded by the surface S we find $\int_V \frac{\partial\rho}{\partial t}d^3x = \frac{d}{dt}\int_V \rho d^3x = -\int_V \vec{\nabla}\cdot\vec{j}d^3x = -\int_S \vec{j}\cdot d\vec{S}$. This implies that probability cannot be created or destroyed; it can only flow from one point to another.

Furthermore, this implies

$$\nabla \cdot (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) = 0$$

$$\Rightarrow \frac{\partial}{\partial t}\left(\frac{1}{c^2}\left(\Psi^* \partial_t \Psi - \Psi \partial_t \Psi^*\right)\right) - \vec{\nabla} \cdot \left(\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*\right) = 0\,. \tag{2.16}$$

After multiplying through by ic^2, in order to make ρ real, we can write this as desired as a continuity equation with

$$\rho = i\left(\Psi^* \partial_t \Psi - \Psi \partial_t \Psi^*\right) \text{ and } \vec{J} = -ic^2\left(\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*\right)\,. \tag{2.17}$$

Now because of the negative sign between the two terms in ρ, the probability density can both take positive and negative values (in contrast to non-relativistic QM where $\rho = |\Psi|^2$ is positive definite)! This is an absurd and nonsensical result for a probability density!!

Junk the KG equation for the moment and try harder. Schrödinger was the first to write down this relativistic wave equation, but discarded it for a different reason; the spectrum is not bounded from below. This was the historical route. See later for a rebirth of the KG equation thanks to Feynman.

2.3 The Dirac Equation

Let us go back to our starting point

$$\widehat{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t} \tag{2.18}$$

and try to give a meaning to the square root in

$$\widehat{H} = \sqrt{m^2c^4 + \widehat{\vec{p}}c^2} = \sqrt{m^2c^4 - \hbar^2 c^2 \vec{\nabla}^2}\,. \tag{2.19}$$

Since the equation is *linear* in $\frac{\partial}{\partial t}$, Lorentz covariance suggests it should be linear also in the $\frac{\partial}{\partial x^i}$, $i = 1, 2, 3$.

So we write

$$\begin{aligned} \widehat{H} &= c\vec{\alpha} \cdot \widehat{\vec{p}} + \beta mc^2 \\ &= -i\hbar c \vec{\alpha} \cdot \vec{\nabla} + \beta mc^2\,, \end{aligned} \tag{2.20}$$

where α^i and β are coefficients to be determined. More explicitly this equation can be written

$$\widehat{H} = -i\hbar c\left(\alpha^1 \frac{\partial}{\partial x^1} + \alpha^2 \frac{\partial}{\partial x^2} + \alpha^3 \frac{\partial}{\partial x^3}\right) + \beta mc^2\,. \tag{2.21}$$

Now we determine the coefficients α^i and β by requiring that this linear operator "squares" to the KG operator

$$\widehat{H}^2 = -\hbar^2 c^2 \vec{\nabla}^2 + m^2 c^4 \,. \tag{2.22}$$

We find

$$\begin{aligned} \widehat{H}^2 &= -\hbar^2 c^2 \left((\alpha^1)^2 \frac{\partial^2}{\partial (x^1)^2} + (\alpha^2)^2 \frac{\partial^2}{\partial (x^2)^2} + (\alpha^3)^2 \frac{\partial^2}{\partial (x^3)^2} \right) + \beta^2 m^2 c^4 \\ & -i\hbar m c^3 \left((\alpha^1 \beta + \beta \alpha^1) \frac{\partial}{\partial x^1} + \ldots \right) \\ & -\hbar^2 c^2 \left((\alpha^1 \alpha^2 + \alpha^2 \alpha^1) \frac{\partial^2}{\partial x^1 \partial x^2} + \ldots \right) . \end{aligned} \tag{2.23}$$

Thus we need to solve

$$\begin{aligned} & (\alpha^i)^2 = \mathbb{I} \,,\; i = 1,2,3 \\ & \alpha^i \alpha^j + \alpha^j \alpha^i = 0 \,,\; i \neq j \\ & \beta^2 = \mathbb{I} \\ & \alpha^i \beta + \beta \alpha^i = 0 \,,\; i = 1,2,3 \,, \end{aligned} \tag{2.24}$$

where $\mathbb{I}$ denotes a unit matrix (if a subscript is added it denotes the dimensionality, e.g. $\mathbb{I}_2$ denotes a 2×2 unit matrix).

It is obviously NOT possible to solve those equations if the coefficients are simply complex numbers. So let us assume that they are $N \times N$ matrices. With some (guess)work it can be shown that the smallest value of N for which eq. (2.24) can be solved is $N = 4$. This implies that the Dirac wave function is a 4-component column vector

$$\Psi(x) = \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \\ \Psi_3(x) \\ \Psi_4(x) \end{pmatrix} \tag{2.25}$$

where $x \equiv (x^0, \vec{x})$ and the Dirac equation becomes a matrix equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \widehat{H} \Psi = (c \vec{\alpha} \cdot \widehat{\vec{p}} + \beta m c^2) \Psi = (-i\hbar c \vec{\alpha} \cdot \vec{\nabla} + \beta m c^2) \Psi \,. \tag{2.26}$$

This is a set of 4 first order linear differential equations to determine $\Psi_1, \ldots, \Psi_4$.

2.4 Representation of the Dirac Matrices

A particular set of solutions of (2.24) for the 4×4 matrices α^i , β can be written with the help of the 2×2 Pauli matrices σ^i , $i = 1,2,3$,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;,\; \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;,\; \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.27}$$

which obey the following identities

$$\begin{aligned} \left(\sigma^i\right)^2 &= \mathbb{I}_2\,,\ i=1,2,3 \\ \sigma^i\sigma^j + \sigma^j\sigma^i &= 0\,,\ i\neq j\,. \end{aligned} \tag{2.28}$$

We may satisfy the first two lines in eq. (2.24) by taking α^i to be the 4×4 matrices

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}\ ,\ \ i=1,2,3\,. \tag{2.29}$$

Now we may satisfy the remaining two lines in eq. (2.24) by taking

$$\beta = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}. \tag{2.30}$$

Because the σ^i are Hermitian, so are the α^i and β, *i.e.*

$$(\alpha^i)^\dagger = \alpha^i\ \ ,\ \ \beta^\dagger = \beta\,. \tag{2.31}$$

($\dagger \leftrightarrow$ complex conjugate transposed)

2.5 Probability Density for the Dirac Equation

The Dirac equation is given by

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar\vec{\alpha}\cdot\vec{\nabla}\Psi + mc^2\beta\Psi \tag{2.32}$$

and its Hermitian conjugate is

$$-i\hbar\frac{\partial\Psi^\dagger}{\partial t} = i\hbar\vec{\nabla}\Psi^\dagger\cdot\vec{\alpha} + mc^2\Psi^\dagger\beta\,. \tag{2.33}$$

(Recall that α^i and β are hermitian and $(AB)^\dagger = B^\dagger A^\dagger$.)

Now take $\Psi^\dagger\times$(Dirac eqn.) and (Hermitian conjugate eqn.)$\times\Psi$ and subtract the two to obtain:

$$i\hbar\left(\Psi^\dagger\frac{\partial\Psi}{\partial t} + \frac{\partial(\Psi^\dagger)}{\partial t}\Psi\right) = -i\hbar c\left(\Psi^\dagger\vec{\alpha}\cdot\vec{\nabla}\Psi + \vec{\nabla}(\Psi^\dagger)\cdot\vec{\alpha}\Psi\right)\,. \tag{2.34}$$

Dividing this equation by $i\hbar$ we obtain a Continuity Equation

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{2.35}$$

with the *positive definite probability density* given by

$$\rho = \Psi^\dagger \Psi = \sum_{k=1}^{4} \Psi_k^* \Psi_k = \sum_{k=1}^{4} |\Psi_k|^2 > 0\,, \tag{2.36}$$

and the probability current

$$\vec{J} = c\Psi^\dagger \vec{\alpha} \Psi\,. \tag{2.37}$$

2.6 Extreme Non-Relativistic Limit of the Dirac Equation

For a particle at rest ($\vec{p} = 0$) the Dirac equation becomes

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= \beta mc^2 \Psi \\ \Rightarrow \frac{\partial \Psi}{\partial t} &= -i\frac{mc^2}{\hbar}\beta\Psi\,. \end{aligned} \tag{2.38}$$

Taking $\beta = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}$ the four equations for the components of Ψ turn into

$$\begin{aligned} \frac{\partial \Psi_1}{\partial t} &= -i\frac{mc^2}{\hbar}\Psi_1 \\ \frac{\partial \Psi_2}{\partial t} &= -i\frac{mc^2}{\hbar}\Psi_2 \\ \frac{\partial \Psi_3}{\partial t} &= +i\frac{mc^2}{\hbar}\Psi_3 \\ \frac{\partial \Psi_4}{\partial t} &= +i\frac{mc^2}{\hbar}\Psi_4 \end{aligned} \tag{2.39}$$

$$\Rightarrow \Psi_1 = c_1 e^{-i\frac{mc^2}{\hbar}t} \text{ e.t.c.} \tag{2.40}$$

where c_1 is an arbitrary constant.

Thus the general solution takes the form

$$\Psi = \begin{pmatrix} c_1 e^{-i\frac{mc^2}{\hbar}t} \\ c_2 e^{-i\frac{mc^2}{\hbar}t} \\ c_3 e^{i\frac{mc^2}{\hbar}t} \\ c_4 e^{i\frac{mc^2}{\hbar}t} \end{pmatrix} \tag{2.41}$$

which can be rewritten as

$$\Psi = e^{-i\frac{mc^2}{\hbar}t} \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix} + e^{i\frac{mc^2}{\hbar}t} \begin{pmatrix} 0 \\ 0 \\ c_3 \\ c_4 \end{pmatrix} \tag{2.42}$$

By acting with the Hamiltonian operator $\widehat{H} = i\hbar\frac{\partial}{\partial t}$ we find that the first term in the solution (2.42) carries *positive energy* $(+mc^2)$ whereas the second term carries *negative energy* $(-mc^2)$.

Although we found a positive probability density (contrary to the KG equation) we find that also the Dirac equation has both positive and negative energy solutions. We shall later interpret the negative energy part as due to *Anti-particles*.

2.7 Spin of the Dirac Particles

The (free) Dirac equation is

$$i\hbar\partial_t\Psi = \widehat{H}\Psi \text{ with } \widehat{H} = c\vec{\alpha}\cdot\widehat{\vec{p}} + \beta mc^2 \tag{2.43}$$

with $\widehat{\vec{p}} = -i\hbar\vec{\nabla}$.

Consider the total angular momentum operator

$$\vec{J} = \vec{L} + \vec{S} = \vec{x}\times\widehat{\vec{p}} + \frac{\hbar}{2}\vec{\Sigma} \text{ where } \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \tag{2.44}$$

Using the Uncertainty Principle

$$[x^i, \widehat{p}^j] = i\hbar\delta_{ij}\,, \tag{2.45}$$

it can be shown with some effort that

$$[\widehat{H}, \vec{S}] = i\hbar c(\vec{\alpha}\times\widehat{\vec{p}}) \tag{2.46}$$

and

$$[\widehat{H}, \vec{L}] = -i\hbar c(\vec{\alpha}\times\widehat{\vec{p}})\,, \tag{2.47}$$

thus, $[\widehat{H}, \widehat{\vec{J}}]$=0. So $\vec{J}$ is a conserved quantity which we interpret as the total angular momentum.

The 3-component of Spin, S_z, is

$$S^3 = \frac{\hbar}{2}\Sigma^3 \tag{2.48}$$

is the matrix $\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ which has eigenvalues $\frac{\hbar}{2}$, $-\frac{\hbar}{2}$, $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$.

$\Rightarrow$ we are describing spin $\frac{1}{2}$ particles (and anti-particles).

2.8 The Covariant Form of the Dirac Equation

Multiply the Dirac equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar c\sum_{i=1}^{3}\alpha^i\frac{\partial\Psi}{\partial x^i} + \beta mc^2\Psi \tag{2.49}$$

with $\frac{\beta}{c}$ to obtain

$$i\hbar\left(\beta\frac{\partial\Psi}{\partial(ct)} + \sum_{i=1}^{3}\beta\alpha^i\frac{\partial\Psi}{\partial x^i}\right) = mc\Psi\,. \tag{2.50}$$

Next define the matrices

$$\gamma^0 = \beta\,,\ \gamma^i = \beta\alpha^i\,,\ i = 1,2,3 \tag{2.51}$$

which are also called *Gamma-Matrices* and allow us to rewrite the Dirac equation as

$$i\hbar\left(\gamma^0\frac{\partial\Psi}{\partial x^0} + \sum_{i=1}^{3}\gamma^i\frac{\partial\Psi}{\partial x^i}\right)\Psi = mc\Psi \tag{2.52}$$

$$\rightarrow i\hbar\gamma\cdot\nabla\Psi = mc\Psi \tag{2.53}$$

where $\nabla^\mu \equiv \left(\frac{\partial}{\partial x^0}, -\vec{\nabla}\right)$ as defined in section 2.1 and

$$\gamma^\mu \equiv \left(\gamma^0, \vec{\gamma}\right) = \left(\gamma^0, \gamma^1, \gamma^2, \gamma^3\right)\,, \tag{2.54}$$

which makes the Dirac equation now look Lorentz covariant.

Dirac or Feynman Slash Notation

For any 4-vector A^μ, we define $\not{A} \equiv \gamma\cdot A = \gamma^0A^0 - \vec{\gamma}\cdot\vec{A}$. Then, the Dirac equation is

$$i\hbar\not{\nabla}\Psi = mc\Psi \tag{2.55}$$

or

$$\not{\hat{p}}\Psi = mc\Psi \tag{2.56}$$

$$\Rightarrow \left(\not{\hat{p}} - mc\right)\Psi = 0 \tag{2.57}$$

where $\hat{p}^\mu = i\hbar\nabla^\mu$.

2.9 Properties of the γ-Matrices

Using the properties of the α^i and β we may show that the gamma-matrices obey the following anti-commutation identity

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\mathbb{I}_4\,, \tag{2.58}$$

with $g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

In particular, $(\gamma^0)^2 = \mathbb{I}_4$, $(\gamma^i)^2 = -\mathbb{I}_4$, $i = 1, 2, 3$.

Furthermore $(\gamma^0) = \beta^\dagger = \beta$ *i.e.* γ^0 is Hermitian, whereas $(\gamma^i)^\dagger = (\beta\alpha^i)^\dagger = (\alpha^i)^\dagger\beta^\dagger = \alpha^i\beta = -\beta\alpha^i = -\gamma^i$ *i.e.* γ^i is anti-Hermitian, $i = 1, 2, 3$.

Using the explicit matrices α^i and β of Section 2.4 we find

$$\gamma^0 = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \ , \ \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} . \tag{2.59}$$

2.10 Digression on Contravariant and Covariant Vectors

A 4-vector, a^μ, may be regarded as a *contravariant vector* (under LT's) and $g_{\mu\nu} \equiv g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ as a Metric Tensor. Then, $a_\mu \equiv \sum_{\nu=0}^{3} g_{\mu\nu}a^\nu \equiv g_{\mu\nu}a^\nu$ is a *covariant vector*. Summation over repeated indices is understood (Einstein summation convention), where one of the indices always is an upper (contravariant) index and the other is a lower (covariant) index.

We may write $a \cdot b = g_{\mu\nu}a^\mu b^\nu = a^\mu b_\mu$. Alternatively, since $g_{\mu\nu} = g_{\nu\mu}$ we can write $a \cdot b = g_{\mu\nu}a^\mu b^\nu = g_{\nu\mu}a^\mu b^\nu = a_\nu b^\nu$

$\rightarrow a \cdot b = a_\mu b^\mu$

In this notation, $\nabla \cdot \nabla = \nabla_\mu \nabla^\mu = \nabla^\mu \nabla_\mu$, the KG equation is

$$\left(\nabla_\mu \nabla^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\phi = 0\,, \tag{2.60}$$

and the Dirac equation is

$$i\hbar\gamma^\mu\nabla_\mu\Psi = mc\Psi \ \text{ or } \ \gamma^\mu\widehat{p}_\mu\Psi = mc\Psi\,. \tag{2.61}$$

2.11 Plane Wave Solutions of the Dirac Equation

From now on we will work in *natural units* $\hbar = c = 1$. We look for plane wave solutions of the Dirac equation of the form

$$\Psi = e^{\mp ip\cdot x} \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.62}$$

where ϕ are the upper two components and χ the lower two components of Ψ and $p \cdot x = Et - \vec{p} \cdot \vec{x}$, with $E > 0$.

The factor $e^{-ip\cdot x}$ gives solutions with positive energy E and momentum $\vec{p}$, and the factor $e^{+ip\cdot x}$ gives solutions with negative energy $-E$ and momentum $-\vec{p}$. Substituting back into the Dirac equation

$$\left(-i\vec{\alpha} \cdot \vec{\nabla} + \beta m\right) \Psi = i\frac{\partial \Psi}{\partial t}\,, \tag{2.63}$$

and using $\partial_t \Psi = \mp iE\Psi$ and $\partial_x \Psi = \pm ip_x \Psi$ e.t.c., we obtain

$$\begin{aligned} (-i(\pm i)\vec{\alpha} \cdot \vec{p} + \beta m)\Psi &= i(\mp iE)\Psi \\ \rightarrow (\pm\vec{\alpha} \cdot \vec{p} + \beta m)\Psi &= \pm E\Psi\,. \end{aligned} \tag{2.64}$$

If we use the standard representation for the α^i and β from Section 2.4 we obtain

$$\begin{pmatrix} m\mathbb{I} & \pm\vec{\sigma} \cdot \vec{p} \\ \pm\vec{\sigma} \cdot \vec{p} & -m\mathbb{I} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \pm E \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.65}$$

which gives the coupled set of equations

$$\begin{aligned} (m \mp E)\phi \pm \vec{\sigma} \cdot \vec{p}\chi &= 0 \\ \pm\vec{\sigma} \cdot \vec{p}\phi - (m \pm E)\chi &= 0\,. \end{aligned} \tag{2.66}$$

We will construct the solutions in such a way that they have a straighforward $\vec{p} = 0$ limit.

Positive Energy Solutions

$$\begin{aligned} (m - E)\phi + \vec{\sigma} \cdot \vec{p}\chi &= 0 \\ \vec{\sigma} \cdot \vec{p}\phi - (m + E)\chi &= 0\,. \end{aligned} \tag{2.67}$$

For $\vec{p} = 0$, $E = m$ and $\chi = 0$ (in agreement with Section 2.6).

For $\vec{p} \neq 0$ it is convenient to solve for χ in terms of ϕ using the second equation in (2.67), *i.e.*

$$\chi = \frac{\vec{\sigma} \cdot \vec{p}}{(E + m)}\phi \tag{2.68}$$

then,

$$\Psi = e^{-ip\cdot x}\begin{pmatrix} \phi \\ \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)}\phi \end{pmatrix} . \tag{2.69}$$

Note that the first equation in (2.67) only gives the on-(mass)shell condition $E^2 = \vec{p}^2 + m^2$.

We can write ϕ in terms of $\phi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\phi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

There are thus two independent positive energy solutions

$$\Psi = e^{-ip\cdot x}U(p,s) \ , \ \ s = 1,2 \tag{2.70}$$

where

$$U(p,s) = \sqrt{E+m}\begin{pmatrix} \phi_s \\ \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)}\phi_s \end{pmatrix} \tag{2.71}$$

is a positive energy Dirac spinor. In the last expression a convenient normalization factor has been introduced.

It can be checked that the first equation in (2.67) is automatically satisfied by using the identity $(\vec{\sigma}\cdot\vec{p})^2 = \vec{p}^2\mathbb{I}_2$.

Negative Energy Solutions

$$\begin{aligned} (m+E)\phi - \vec{\sigma}\cdot\vec{p}\chi &= 0 \\ -\vec{\sigma}\cdot\vec{p}\phi - (m-E)\chi &= 0\,. \end{aligned} \tag{2.72}$$

For $\vec{p} = 0$, $E = m$ and $\phi = 0$ (in agreement with Section 2.6).

For $\vec{p} \neq 0$ it is convenient to solve for ϕ in terms of χ using the first equation in (2.72), *i.e.*

$$\phi = \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)}\chi \tag{2.73}$$

then,

$$\Psi = e^{+ip\cdot x}\begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)}\chi \\ \chi \end{pmatrix} . \tag{2.74}$$

Note that the second equation in (2.72) only gives the on-(mass)shell condition $E^2 = \vec{p}^2 + m^2$.

We can write χ in terms of $\chi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\chi_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

There are thus two independent negative energy solutions

$$\Psi = e^{+ip\cdot x} V(p,s) \ , \ \ s = 1,2 \tag{2.75}$$

where

$$V(p,s) = \sqrt{E+m} \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)}\chi_s \\ \chi_s \end{pmatrix} \tag{2.76}$$

is a negative energy Dirac spinor. In the last expression a convenient normalization factor has been introduced.

It can be checked that the second equation in (2.72) is automatically satisfied by using the identity $(\vec{\sigma}\cdot\vec{p})^2 = \vec{p}^2\mathbb{I}_2$.

Interpretation

To find the physical interpretation for the four independent solutions we consider the rest frame $\vec{p} = 0$. Then:

$$\begin{aligned}
U(p,1) &= \sqrt{2m} \begin{pmatrix} \phi^1 \\ 0 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
U(p,2) &= \sqrt{2m} \begin{pmatrix} \phi^2 \\ 0 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\
V(p,1) &= \sqrt{2m} \begin{pmatrix} 0 \\ \chi^1 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\
V(p,2) &= \sqrt{2m} \begin{pmatrix} 0 \\ \chi^2 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}
\end{aligned} \tag{2.77}$$

Furthermore, for $\vec{p} = 0$ we have $\vec{L} = \vec{x} \times \vec{p} = 0$ so that the total angular momentum operator becomes

$$\begin{aligned}
\vec{J} &= \vec{L} + \vec{S} = \vec{S} = \frac{1}{2}\vec{\Sigma} \\
&\Rightarrow \ S_z = \begin{pmatrix} \frac{1}{2}\sigma_3 & 0 \\ 0 & \frac{1}{2}\sigma_3 \end{pmatrix}
\end{aligned} \tag{2.78}$$

Thus, $U(p,1)$ and $U(p,2)$ are *positive energy solutions* with S_z eigenvalues $s_z = +1/2$ and $s_z = -1/2$ respectively, whereas $V(p,1)$ and $V(p,2)$ are *negative energy solutions* with S_z eigenvalues $s_z = -1/2$ and $s_z = +1/2$ respectively.

In general, U and V are the Lorentz boosts of these solutions to a frame where $\vec{p} \neq 0$. Interpret the negative energy solutions later.

2.12 Properties of Solutions

Since $e^{-ip\cdot x}U(p,s)$ is a solution of the Dirac equation

$$\begin{aligned} & (i\gamma\cdot\nabla - m)(e^{-p\cdot x}U(p,s)) = 0 \\ \rightarrow \quad & \left(i\gamma^0\frac{\partial}{\partial t} + i\vec{\gamma}\cdot\vec{\nabla} - m\right)e^{-i(Et-\vec{p}\cdot\vec{x})}U(p,s) = 0 \\ \rightarrow \quad & (E\gamma^0 - \vec{p}\cdot\vec{\gamma} - m)U = 0 \\ \rightarrow \quad & (p\cdot\gamma - m)U = 0 \\ \Rightarrow \quad & (\not{p} - m)U(p,s) = 0 \end{aligned} \tag{2.79}$$

i.e. $U(p,s)$ obeys the Dirac equations with $\widehat{p}^\mu$ simply replaced by p^μ. Similarly, we find

$$(\not{p} + m)V(p,s) = 0\,. \tag{2.80}$$

We will often make use of the *adjoint spinors* which are defined as

$$\overline{U}(p,s) \equiv U^\dagger(p,s)\gamma^0\,,\;\; \overline{V}(p,s) \equiv V^\dagger(p,s)\gamma^0\,. \tag{2.81}$$

By taking the Hermitian adjoint of the equation obeyed by U and V we find

$$\overline{U}(p,s)(\not{p} - m) = 0\,,\;\; \overline{V}(p,s)(\not{p} + m) = 0\,. \tag{2.82}$$

One may also check directly (using again $(\vec{\sigma}\cdot\vec{p})^2 = \vec{p^2}\mathbb{I}_2$) that

$$U^\dagger(p,s)U(p,s) = V^\dagger(p,s)V(p,s) = 2E\,,\;\; s = 1,2\,, \tag{2.83}$$

and

$$\overline{U}(p,s)U(p,s) = 2m\,,\;\; \overline{V}(p,s)V(p,s) = -2m\,,\;\; s = 1,2\,. \tag{2.84}$$

2.13 Anti-Particles — Hole Theory

Since the Dirac equation has negative energy solutions, why do positive energy electrons not radiate energy and fall into a negative energy state? Dirac: Negative energy states are completely *filled* and the Pauli exclusion principle (which applies to fermions) forbids the transition. Consequently, the Vacuum is a state with all positive energy states empty *but* all negative energy states filled.

(picture goes here)

If a photon excites a negative energy e^- of energy $-|E_2|$ into a positive energy e^- of energy $|E_1|$, we observe the production of an e^- of mass m, charge $-|e|$ and energy $|E_1|$, and a *Hole* in the negative energy sea (Pair production). Note that there is a *gap* of $2mc^2$ between the negative and positive energy states and, hence, the photon energy $h\nu = |E_1| + |E_2|$ must be larger than $2mc^2$ for this to happen. The hole appears as a particle of mass m, charge $+|e|$ and energy $+|E_2|$.
$\Rightarrow$ The existence of the Positron (and Anti-particles in general) is predicted!

(picture goes here)

The absence of a spin-up electron of energy $-|E|$ and momentum $-\vec{p}$ is equivalent to the presence of a spin-down positron of energy $+|E|$ and momentum $+\vec{p}$. (Think about time running backwards or the arrow in a Feynman diagram reversed)

(picture goes here)

Thus, the electron wavefunction $e^{ip\cdot x}V(p,s)$ corresponding to energy $-E$ and momentum $-\vec{p}$ describes a positron of energy $+E$ and momentum $+\vec{p}$. Also, $V(p,1)$ and $V(p,2)$ which describe spin down and spin up negative energy electrons must describe spin up and spin down positrons.

2.14 Vacuum Polarization

In general the infinite negative charge of the vacuum produces no effect because the distribution of charge is homogeneous.

However, consider the effect of a positive energy electron with charge $-|e|$ on the vacuum. It repels the negative energy electrons and electrically polarises the vacuum. Thus the physical charge $-|e|$ seen by a test charge at a large distance from the electron is numerically smaller than the bare charge $-|e_0|$, i.e. $|e| < |e_0|$.

(picture goes here)

However, if the test charge comes very close it will see the bare charge $-|e_0|$. For S-wave electrons ($l = 0$) in an atom, the proton sees a charge numerically greater than the ordinary electric charge $|e|$. Note that for $l > 0$ the wavefunction vanishes at the origin and the proton feels a numerically smaller charge. This effect leads to measurable shifts of the energy levels of atoms.

2.15 Charge Conjugation Symmetry $\mathcal{C}$

We construct an operator acting on the Dirac wave function

$$\mathcal{C} : \Psi \to \Psi_C \tag{2.85}$$

which turns a positive energy electron wavefunction (e^-) into a negative energy wave function (e^+) with the same momentum and spin state. If $\Psi = e^{-ip\cdot x}U(p,s)$ then $\Psi_C = e^{ip\cdot x}V(p,s)$. The required operation turns out to be

$$\mathcal{C} : \Psi \to \Psi_C = C\gamma^0\Psi^*\,, \tag{2.86}$$

where $C = i\gamma^2\gamma^0$. Useful properties: $C^\dagger = -C$, $C^2 = -\mathbb{I}$, $C^{-1} = -C$ and $C\gamma^\mu C = (\gamma^\mu)^T$.

A symmetry of a wave equation is an operation on a wave function $\Psi \to \Psi'$ and on the space-time coordinates $x \to x'$ such that Ψ' obeys the same equation as Ψ, with x replaced by x'.

$\Psi \to \Psi_C$, $x \to x'$ can be shown to be a symmetry of the Dirac equation as follows:

We say that the Dirac equation is charge conjugation invariant. The Dirac equation may be written as

$$(i\gamma^\mu\nabla_\mu - m)\Psi = 0, \tag{2.87}$$

taking the complex conjugate gives

$$(-i(\gamma^\mu)^*\nabla_\mu - m)\Psi^* = 0\,. \tag{2.88}$$

Now multiply from the left with $C\gamma^0$

$$\begin{aligned} & (-iC\gamma^0(\gamma^\mu)^*\nabla_\mu - m)\Psi^* = 0 \\ \to\ & (i\gamma^\mu(C\gamma^0))\nabla_\mu - mc\gamma^0)\Psi^* = 0 \\ \to\ & (i\gamma^\mu\nabla_\mu - m)\Psi_C = 0\,, \end{aligned} \tag{2.89}$$

where we have used the identity $C\gamma^0(\gamma^\mu)^* = -\gamma^\mu(C\gamma^0)$. This shows that Ψ_C obeys the same equation as Ψ.

2.16 Space Inversion $\mathcal{P}$

The Dirac equation is also invariant under reflection of space coordinates in the origin

$$\mathcal{P} : \vec{x} \to \vec{x}' = -\vec{x}\,,\ t \to t' = t\,. \tag{2.90}$$

The corresponding operation[7] on Dirac spinors is

$$\mathcal{P} : \Psi \to \Psi' = P\Psi \tag{2.91}$$

with $P = \gamma^0$. It can be checked by direct calculation that $PU(\vec{p}, s) = U(-\vec{p}, s)$ i.e. $\vec{p} \to -\vec{p}$ as expected for space inversion, but the spin state is unchanged. Also $PV(\vec{p}, s) = -V(-\vec{p}, s)$. The -1 factor indicates that anti-particles have opposite Parity to particles.

In this case, to check invariance of the Dirac equation, it is necessary to replace ∂_x, ∂_y and ∂_z by $-\partial_x$, $-\partial_y$ and $-\partial_z$, as well as replacing Ψ by Ψ', i.e. Ψ' obeys the same equation as Ψ with ∂_x, ∂_y and ∂_z replaced by $-\partial_x$, $-\partial_y$ and $-\partial_z$.

2.17 Time Reversal $\mathcal{T}$

The Dirac equation also has a symmetry under time reversal

$$\mathcal{T} : t \to t' = -t\,,\ \vec{x} \to \vec{x}' = \vec{x}\,, \tag{2.92}$$

the appropriate transformation of Ψ is

$$\mathcal{T} : \Psi \to \Psi' = T\psi^* \tag{2.93}$$

with $T = -\gamma^1\gamma^3$. It can be checked directly that this is the correct transformation by showing that

$$TU^*(\vec{p}, 1) = +U(-\vec{p}, 2)\,,\ TV^*(\vec{p}, 1) = -V(-\vec{p}, 2) \tag{2.94}$$

Thus, the transformation changes a solution of the Dirac equation with momentum $\vec{p}$ and spin up into a solution with momentum $-\vec{p}$ and spin down.

This is as expected for time reversal, since $\vec{p} = m\vec{v}/\sqrt{1 - \vec{v}^2/c^2}$ and $\vec{L} = \vec{x} \times \vec{p}$, and thus under time reversal $\vec{p} \to -\vec{p}$ and $\vec{L} \to -\vec{L}$ and in particular $L_z \to -L_z$. We assume that this applies to any AM operator, so that in particular $S_z \to -S_z$. In this case, Ψ' obeys the same equation as Ψ with ∂_t replaced by $-\partial_t$.

2.18 Dirac Covariants

It is important in the study of the Weak Interactions to know the properties of objects like $\overline{\Psi}\gamma^\mu\Psi$, $\overline{\Psi}\gamma^\mu\gamma_5\Psi$, etc, where we introduced

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}\,. \tag{2.95}$$

[7]Since $\mathcal{P}$ also transforms the space time coordinates this operation should be written more properly as $\Psi(t, \vec{x}) \to \Psi'(t', \vec{x}') = \Psi'(t, -\vec{x}) = P\Psi(t, \vec{x})$. Hence $\Psi'(t, \vec{x}) = P\Psi(t, -\vec{x})$; a similar comment applies to time reversal T.

We list here the behaviour of some of the Dirac covariants under Lorentz transformations, $\mathcal{P}$, $\mathcal{C}$:

Covariant	LT's	$\mathcal{P}$	$\mathcal{C}$
$\overline{\Psi}\Psi$	scalar	$+\overline{\Psi}\Psi$	$-\overline{\Psi}\Psi$
$\overline{\Psi}\gamma_5\Psi$	pseudoscalar	$-\overline{\Psi}\gamma_5\Psi$	$-\overline{\Psi}\gamma_5\Psi$
$\overline{\Psi}\gamma^\mu\Psi$	4-vector	$+\overline{\Psi}\gamma^0\Psi$ $-\overline{\Psi}\gamma^i\Psi$	$+\overline{\Psi}\gamma^\mu\Psi$
$\overline{\Psi}\gamma^\mu\gamma_5\Psi$	(pseudo) 4-vector	$-\overline{\Psi}\gamma^0\gamma_5\Psi$ $+\overline{\Psi}\gamma^i\gamma_5\Psi$	$-\overline{\Psi}\gamma^\mu\gamma_5\Psi$

(γ_5 has the properties $\{\gamma_5, \gamma^\mu\} = 0$, $\gamma_5^\dagger = \gamma_5$)

For example under P the behaviour of the vector current is

$$\begin{aligned}\overline{\Psi}\gamma^\mu\Psi \;\;&\rightarrow\;\; \overline{\Psi}'\gamma^\mu\Psi' \\ &= (\Psi')^\dagger\gamma^0\gamma^\mu\Psi' \\ &= \Psi^\dagger(\gamma^0)^\dagger\gamma^0\gamma^\mu\gamma^0\Psi \\ &= \Psi^\dagger\gamma^0\gamma^0\gamma^\mu\gamma^0\Psi \\ &= \overline{\Psi}\gamma^0\gamma^\mu\gamma^0\Psi \end{aligned} \tag{2.96}$$

Thus,

$$\overline{\Psi}\gamma^0\Psi \rightarrow \overline{\Psi}\gamma^0\gamma^0\gamma^0\Psi = \overline{\Psi}\gamma^0\Psi \tag{2.97}$$

$$\begin{aligned}\overline{\Psi}\gamma^i\Psi \;\;&\rightarrow\;\; \overline{\Psi}\gamma^0\gamma^i\gamma^0\Psi \\ &= -\overline{\Psi}(\gamma^0)^2\gamma^i\Psi \\ &= -\overline{\Psi}\gamma^i\Psi \end{aligned} \tag{2.98}$$

In the Relativistic version of Time Dependent Perturbation Theory (Feynman Diagrams) the probability amplitudes for Electromagnetic Scattering of 2 particles via photon exchange contains a factor

$$\begin{aligned}&\overline{U}(p_2, s_2)\gamma_\mu U(p_1, s_1)\overline{U}(p_4, s_4)\gamma^\mu U(p_3, s_3) \\ &= \overline{U}\gamma^0 U\overline{U}\gamma^0 U - \sum_{i=1}^{3}\overline{U}\gamma^i U\overline{U}\gamma^i U\,, \end{aligned} \tag{2.99}$$

which is *invariant under both $\mathcal{P}$ and $\mathcal{C}$* because the two negative signs for the $\overline{U}\gamma^i U$ cancel for space inversion. Thus the Electromagnetic interactions are both space reflection invariant and charge conjugation invariant.

(picture goes here)

The corresponding probability amplitude for the Weak Interaction has a factor

$$\begin{aligned}&\overline{U}(p_2,s_2)\gamma_\mu(\mathbb{I}-\gamma_5)U(p_1,s_1)\overline{U}(p_4,s_4)\gamma^\mu(\mathbb{I}-\gamma_5)U(p_3,s_3)\\&=\overline{U}\gamma_\mu U\overline{U}\gamma^\mu U+\overline{U}\gamma\mu\gamma_5 U\overline{U}\gamma^\mu\gamma_5 U\\&-\overline{U}\gamma_\mu\gamma_5 U\overline{U}\gamma^\mu U-\overline{U}\gamma_\mu U\overline{U}\gamma^\mu\gamma_5 U\,.\end{aligned}\tag{2.100}$$

The term

$$\begin{aligned}&\overline{U}\gamma_\mu\gamma_5 U\overline{U}\gamma^\mu U\\&=\overline{U}\gamma^0\gamma_5 U\overline{U}\gamma^0 U-\sum_{i=1}^{3}\overline{U}\gamma^i\gamma_5 U\overline{U}\gamma^i U\end{aligned}\tag{2.101}$$

changes sign under both $\mathcal{P}$ and $\mathcal{C}$ transformations, because $\overline{\Psi}\gamma^\mu\Psi$ and $\overline{\Psi}\gamma^\mu\gamma_5\Psi$ transform with opposite signs both for $\mu=0$ and $\mu=i$. Thus, the weak interactions *break both space inversion and charge conjugation invariance*. This manifests itself in the angular dependence of scattering processes (e.g. $\cos\theta$ changes sign under space inversion: $0\to\pi-0$). Note however that the combined action of $\mathcal{C}$ and $\mathcal{P}$, $\mathcal{CP}$, is a symmetry of this interaction.

Note, that in a general theory $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$ are not preserved, but the combination of the three transformations $\mathcal{CPT}$ is always a symmetry.

2.19 Neutrinos

Some modification of RQM is needed in the physically important case of massless spin-1/2 particles — Neutrinos (with todays experimental evidence of Neutrino oscillations this is not quite true, nevertheless it is a very good approximation.)

First, define the Helicity of a particle as the component of its AM $\vec{J}=\vec{L}+\vec{\Sigma}/2$ in its direction of motion. For a Dirac particle,

$$\text{Helicity}=\vec{J}\cdot\frac{\vec{p}}{|\vec{p}|}=\frac{\vec{\Sigma}}{2}\cdot\frac{\vec{p}}{|\vec{p}|}\tag{2.102}$$

because $\vec{L}\cdot\vec{p}=(\vec{x}\times\vec{p})\cdot\vec{p}=0$.

Experimental observation shows that whereas an e^- can have Helicity $+1/2$ or $-1/2$, a Neutrino (which is massless) can only have Helicity $-1/2$ and an Anti-Neutrino can only have Helicity $+1/2$.

Thus, whereas we need four degrees of freedom to describe the 2 spin states of an electron or positron, we need only 2 degrees of freedom to describe the spin states of the neutrino and anti-neutrino. We need to discard 2 spin states of the Dirac particle.

Now we return to the Dirac equation for a positive energy solution of energy E and momentum $\vec{p}$. However, we choose a different representation of the Dirac matrices (and hence a different representation of the gamma-matrices). This does not effect the physics but makes the proof much easier. It may be checked that

$$\beta = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix} \tag{2.103}$$

also obey the Dirac Algebra (2.24).

For $\Psi = e^{-ip\cdot x}\begin{pmatrix} \phi \\ \chi \end{pmatrix}$ the Dirac equation reduces to $(\vec{\alpha}\cdot\vec{p}+\beta m)\Psi = E\Psi$ (see Section 2.11), from which we get

$$\begin{pmatrix} \vec{\sigma}\cdot\vec{p} & m\mathbb{I} \\ m\mathbb{I} & -\vec{\sigma}\cdot\vec{p} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix} = E\begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{2.104}$$

which gives the coupled equations

$$\begin{aligned} \vec{\sigma}\cdot\vec{p}\phi + m\chi &= E\phi \\ m\phi - \vec{\sigma}\cdot\vec{p}\chi &= E\chi\,. \end{aligned} \tag{2.105}$$

Now taking $m = 0$ for a massless neutrino decouples the two equations,

$$\begin{aligned} \vec{\sigma}\cdot\vec{p}\phi &= E\phi \\ \vec{\sigma}\cdot\vec{p}\chi &= -E\chi\,, \end{aligned} \tag{2.106}$$

and since $E = |\vec{p}|$ for $m = 0$,

$$\begin{aligned} \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\phi &= \frac{1}{2}\phi \\ \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\chi &= -\frac{1}{2}\chi\,. \end{aligned} \tag{2.107}$$

Thus the upper 2 components of Ψ describe Helicity $1/2$, and the lower two describe helicity $-1/2$ when Ψ has positive energy. To obtain an appropriate Ψ to describe a Neutrino we perform a projection that removes the upper 2 components.

This may be achieved by using $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. With the above choice of $\vec{\alpha}$ and β,

$$\gamma^0 = \beta = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \quad \gamma^i = \beta\alpha^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \tag{2.108}$$

and

$$\gamma^5 = i\begin{pmatrix} -\sigma^1\sigma^2\sigma^3 & 0 \\ 0 & \sigma^1\sigma^2\sigma^3 \end{pmatrix} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \tag{2.109}$$

If we form $\frac{1}{2}(\mathbb{I}-\gamma_5) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I} \end{pmatrix}$, then we may use it to project the upper 2 components and leave the Helicity $-1/2$ components. Thus,

$$\Psi_L \equiv \frac{1}{2}(\mathbb{I} - \gamma_5)\Psi\,, \tag{2.110}$$

with Ψ a positive energy spinor, may be used to describe the neutrino.

If instead we start from a negative energy solution Ψ, the from Section 2.11

$$(-\vec{\alpha}\cdot\vec{p} + \beta m)\Psi = -E\Psi\,. \tag{2.111}$$

For $\Psi = e^{ip\cdot x}\begin{pmatrix} \phi \\ \chi \end{pmatrix}$ we then have

$$\begin{pmatrix} -\vec{\sigma}\cdot\vec{p} & m\mathbb{I} \\ m\mathbb{I} & \vec{\sigma}\cdot\vec{p} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix} = -E\begin{pmatrix} \phi \\ \chi \end{pmatrix}\,, \tag{2.112}$$

which gives the coupled equations

$$\begin{aligned} -\vec{\sigma}\cdot\vec{p}\phi + m\chi &= -E\phi \\ m\phi + \vec{\sigma}\cdot\vec{p}\chi &= -E\chi\,. \end{aligned} \tag{2.113}$$

Now taking $m = 0$ for a massless anti-neutrino

$$\begin{aligned} \vec{\sigma}\cdot\vec{p}\phi &= E\phi \\ \vec{\sigma}\cdot\vec{p}\chi &= -E\chi\,, \end{aligned} \tag{2.114}$$

and since $E = |\vec{p}|$ for $m = 0$,

$$\begin{aligned} \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\phi &= \frac{1}{2}\phi \\ \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\chi &= -\frac{1}{2}\chi\,. \end{aligned} \tag{2.115}$$

This is the same as for the positive energy solution. Thus, the upper 2 components of Ψ still describe Helicity $+1/2$ and the lower 2 components describe Helicity $-1/2$. To obtain an appropriate Ψ to describe an Anti-Neutrino with Helicity $+1/2$ we need a negative energy state with Helicity $-1/2$.

Thus, $\Psi_L = \frac{1}{2}(\mathbb{I} - \gamma_5)\Psi$ with Ψ a negative energy spinor, may be used to describe the anti-neutrino with Helicity $+1/2$.

2.20 Feynman's Interpretation of the Klein-Gordon Equation

In Section 1.2 we abandoned the KG equation because the Probability density

$$\rho = i\left(\phi^* \partial_t \phi - \phi \partial_t \phi^*\right) \tag{2.116}$$

could give negative values (we have renamed the wavefuntion Ψ by ϕ).

It can be checked by direct substitution that the KG equation

$$(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2)\phi = 0 \tag{2.117}$$

has positive energy solutions

$$\phi = Ne^{-ip\cdot x} = Ne^{-i(Et - \vec{p}\cdot\vec{x})} \tag{2.118}$$

and negative energy solutions

$$\phi = Ne^{ip\cdot x} = Ne^{i(Et - \vec{p}\cdot\vec{x})}\,. \tag{2.119}$$

The probability density of such solutions is

$$\rho = |N|^2(\pm 2E)\,. \tag{2.120}$$

Thus, negative probabilities come from negative energy solutions. These are (as usual) the problem.

We need an interpretation for the negative energy solutions of the KG equation. Dirac Hole theory will NOT work for the spin-0 Bosons described by the KG equation, because they do not obey the Dirac exclusion principle to give a filled negative energy sea.

Feynman gave an alternative way of interpreting negative energy solutions which works for both bosons and fermions!

The emission/absorption of an anti-particle with 4-momentum p^μ is equivalent to the absorption/emission of a negative energy particle with 4-momentum $-p^\mu$.

In Feynman diagrams, which are the rules of calculating scattering and decay amplitudes in RQM, when Anti-Particles are involved we draw lines for negative energy particles propagating backwards in time and use Feynman's interpretation. E.g. for electromagnetic Electron-Positron scattering via photon exchange there are two diagrams that contribute:

(picture goes here)

2.21 Dirac Equation in an Electromagnetic Field

In classical relativistic mechanics the interaction of a particle carrying charge q in an external electromagnetic field can be obtained by substituting the momentum as

$$p^\mu \to p^\mu + qA^\mu \,, \tag{2.121}$$

where A^μ is the 4-vector potential

$$A^\mu \equiv (A^0, \vec{A}) = (\phi, \vec{A}) \tag{2.122}$$

with ϕ the scalar potential and $\vec{A}$ the vector potential. (Remember: $\vec{E} = -\vec{\nabla}\phi - \partial_t \vec{A}$, $\vec{B} = \vec{\nabla} \times \vec{A}$).

This works also for RQM

$$\widehat{p}^\mu \to \widehat{p}^\mu + qA^\mu \tag{2.123}$$

or equivalently

$$\nabla^\mu \to \nabla^\mu - iqA^\mu \,. \tag{2.124}$$

The free particle Dirac equation is $(i\gamma \cdot \nabla - m)\Psi = 0$. Making the above substitution

$$\gamma \cdot \nabla \to \gamma \cdot \nabla - iq\gamma \cdot A \tag{2.125}$$

the Dirac equation in an electromagnetic field is

$$(i\gamma \cdot \nabla - m)\Psi = -q\gamma \cdot A \tag{2.126}$$

or

$$(i\not{\nabla} - m)\Psi = -q\not{A} \,. \tag{2.127}$$

It is sometimes convenient to write the equation in terms of the Dirac matrices i.e. in Hamiltonian form. Begin with the equation

$$(i\gamma^0 \frac{\partial}{\partial t} + i\vec{\gamma} \cdot \vec{\nabla} - m)\Psi = -q(A^0\gamma^0 - \vec{A} \cdot \vec{\gamma})\Psi \tag{2.128}$$

and multiply from left with $\gamma^0 = \beta$. Since $(\gamma^0)^2 = \mathbb{I}$ and $\gamma^0\gamma^i = \beta\beta\alpha^i = \alpha^i$ we obtain

$$\begin{aligned} i\frac{\partial \Psi}{\partial t} &= \left((-i\vec{\nabla} + q\vec{A}) \cdot \alpha + \beta m\right)\Psi - qA^0\Psi \\ \to \quad i\frac{\partial \Psi}{\partial t} &= \left(\alpha \cdot \widehat{\vec{\Pi}} + \beta m\right)\Psi - qA^0\Psi \,, \end{aligned} \tag{2.129}$$

where

$$\widehat{\vec{\Pi}} = -i\vec{\nabla} + q\vec{A} = \widehat{\vec{p}} + q\vec{A} \,. \tag{2.130}$$

2.22 The Magnetic Moment of the Electron

In the non-relativistic limit the rest mass mc^2 is the largest energy in the problem (since $|\vec{v}|^2 << c^2$) and we can write for a positive energy solution

$$\Psi = e^{-imt} \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.131}$$

where ϕ and χ vary slowly with time and will be called large and small components for reasons that will become clear in a moment.

Substituting in the Dirac equation (with the Dirac representation for β and α^i, which is more appropriat for studying non-relativistic limits) in an electromagnetic field

$$me^{-imt} \begin{pmatrix} \phi \\ \chi \end{pmatrix} + ie^{-imt} \begin{pmatrix} \partial_t \phi \\ \partial_t \chi \end{pmatrix} = e^{-imt} \begin{pmatrix} (-qA^0 + m)\mathbb{I} & \vec{\sigma} \cdot \widehat{\vec{\Pi}} \\ \vec{\sigma} \cdot \widehat{\vec{\Pi}} & (-qA^0 - m)\mathbb{I} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{2.132}$$

multiplying with e^{+imt} and subtracting the first term on the left hand side from both sides we obtain

$$\begin{pmatrix} i\partial_t \phi \\ i\partial_t \chi \end{pmatrix} = \begin{pmatrix} -qA^0\mathbb{I} & \vec{\sigma} \cdot \widehat{\vec{\Pi}} \\ \vec{\sigma} \cdot \widehat{\vec{\Pi}} & (-qA^0 - 2m)\mathbb{I} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix}. \tag{2.133}$$

The lower eqution is

$$i\partial_t \chi = \vec{\sigma} \cdot \widehat{\vec{\Pi}} \phi - (qA^0 + 2m)\chi \,. \tag{2.134}$$

For χ varying slowly with time and under the assumption $2m >> qA_0$

$$\chi \sim \frac{\vec{\sigma} \cdot \widehat{\vec{\Pi}}}{2m} \phi \tag{2.135}$$

where $\widehat{\vec{\Pi}} = \widehat{\vec{p}} + q\vec{A}$. Hence, for momenta and EM fields small compared to the rest mass

$$\chi << \phi \,. \tag{2.136}$$

Now using the top equation

$$i\partial_t \phi = -qA^0 \phi + \vec{\sigma} \cdot \widehat{\vec{\Pi}} \chi \,, \tag{2.137}$$

we get, using eqn. (2.135)

$$\Rightarrow i\partial_t \phi = -qA^0 \phi + \frac{(\vec{\sigma} \cdot \widehat{\vec{\Pi}})^2}{2m} \phi \,. \tag{2.138}$$

To simplify this further we may use the identity

$$(\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) = \vec{a}\cdot\vec{b}\mathbb{I} + i\vec{\sigma}\cdot(\vec{a}\times\vec{b})\,, \tag{2.139}$$

which follows from

$$\sigma^i\sigma^j = \delta^{ij}\mathbb{I} + i\epsilon^{ijk}\sigma^k \tag{2.140}$$

where summation over k is understood. The Kronecker delta is defined as $\delta^{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$; and $\epsilon^{123} = \epsilon^{231} = \epsilon^{312} = +1$, $\epsilon^{132} = \epsilon^{213} = \epsilon^{321} = -1$ and otherwise $\epsilon^{ijk} = 0$. In terms of the ϵ-tensor

$$\left(\vec{a}\times\vec{b}\right)^i = \epsilon^{ijk}a^jb^k \tag{2.141}$$

and, hence,

$$(\vec{\sigma}\cdot\widehat{\vec{\Pi}})^2\phi = (\vec{\sigma}\cdot\widehat{\vec{\Pi}})(\vec{\sigma}\cdot\widehat{\vec{\Pi}})\phi = \widehat{\vec{\Pi}}\cdot\widehat{\vec{\Pi}}\mathbb{I}\phi + i\vec{\sigma}\cdot(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi\,. \tag{2.142}$$

Now ($\hbar = 1$),

$$(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi = (\widehat{\vec{p}}+q\vec{A})\times(\widehat{\vec{p}}+q\vec{A})\phi = (-i\vec{\nabla}+q\vec{A})\times(-i\vec{\nabla}+q\vec{A})\phi = -(\vec{\nabla}+iq\vec{A})\times(\vec{\nabla}+iq\vec{A})\phi\,. \tag{2.143}$$

The x component of this expression is

$$\begin{aligned} & -(\partial_y + iqA_y)(\partial_z + iqA_z)\phi + (\partial_z + iqA_z)(\partial_y + iqA_y)\phi \\ = \; & -\left[\partial_y(iqA_z\phi) - \partial_z(iqA_y\phi) + iqA_y\partial_z\phi - iqA_z\partial_y\phi\right] \\ = \; & -iq[\partial_yA_z - \partial_zA_y]\phi \\ = \; & -iqB_x\phi\,, \end{aligned} \tag{2.144}$$

and hence

$$(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi = -iq\vec{B}\phi\,. \tag{2.145}$$

Now

$$(\vec{\sigma}\cdot\widehat{\vec{\Pi}})^2\phi = \widehat{\vec{\Pi}}\cdot\widehat{\vec{\Pi}}\phi + q\vec{\sigma}\cdot\vec{B}\phi\,, \tag{2.146}$$

hence, the non-relativistic limit of the Dirac equation in an EM field (also called Pauli equation) may be written as

$$i\frac{\partial\phi}{\partial t} = \left(-qA^0 + \frac{(\widehat{\vec{p}}+q\vec{A})^2}{2m} + \frac{q\vec{\sigma}\cdot\vec{B}}{2m}\right)\phi \tag{2.147}$$

or writing $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$ for the spin of the electron (setting $\hbar = 1$) we obtain

$$i\frac{\partial\phi}{\partial t} = \left(-qA^0 + \frac{(\widehat{\vec{p}}+q\vec{A})^2}{2m} + \frac{q\vec{S}\cdot\vec{B}}{m}\right)\phi \tag{2.148}$$

where ϕ is a two-component wave function for the non-relativistic spin-1/2 particle.

Comparing with the usual form of the non-rel. Schrödinger equation

$$i\frac{\partial\Psi}{\partial t} = \left(-\frac{1}{2m}\vec{\nabla}^2 + V\right)\Psi\,, \tag{2.149}$$

we can interpret the last term in eqn. (2.148) as a potential energy $-\vec{\mu}_{spin}\cdot\vec{B}$ due to the spin magnetic moment of the electron in an external magnetic field. Thus the spin magnetic moment is

$$\vec{\mu}_{spin} = -\frac{q\vec{S}}{m} \equiv -g\frac{q\vec{S}}{2m} \tag{2.150}$$

where g is the so-called gyromagnetic ration. Hence, the Dirac equation predicts $g = 2$ whereas classically we would expect $g = 1$. This prediction was confirmed experimentally and is one of the spectacular successes of the Dirac equation! Including radiative correction from Quantum Electrodynamics (QED) yields a more precise value of $g = 2(1.0011\ldots)$ which agrees up to nine digits after the dot with experiment!

2.23 Hydrogen Atom Spectrum

In the presence of an electrostatic potential $V(r)$ the Dirac equation becomes

$$\widehat{H}\Psi = (\vec{\alpha}\cdot\widehat{\vec{p}} + \beta m + V(r))\Psi = i\frac{\partial\Psi}{\partial t}\,, \tag{2.151}$$

for positive energy solutions with energy eigenvalue $E > 0$ we make an separation ansatz

$$\Psi = e^{-iEt}\Psi_0(r,\theta,\phi) \tag{2.152}$$

so that

$$i\partial_t\Psi = E\Psi \tag{2.153}$$

which gives a time independent equation

$$(\vec{\alpha}\cdot\widehat{\vec{p}} + \beta m + V(r))\Psi = E\Psi\,. \tag{2.154}$$

For a Hydrogen-like atom we take

$$V(r) = -\frac{Z\alpha}{r}\ \ ,\ \ \alpha = \frac{e^2}{4\pi}\,. \tag{2.155}$$

The total AM operator $\vec{J}$ commutes with $(\vec{\alpha}\cdot\widehat{\vec{p}} + \beta m)$ as in section 2.7. Also $\vec{J}$ commutes with $V(r)$ because $V(r)$ is independent of Spin and as in section 1.3, orbital AM operator $\vec{L}$ commutes with $V(r)$. Thus $[\vec{J},\widehat{H}] = 0$.

The problem can be solved using simultaneous eigenstates $\psi^l_{j,m}$ of $\vec{J}^2$, J_z and the parity operator P (which takes $\vec{x}\to-\vec{x}$). The corresponding quantum numbers are $j(j+1)$, m and $(-1)^l$, where l is orbital angular momentum.

For the spin-1/2 electron, the allowed values of j are $j = l \pm 1/2$. The gory details of the calculation can be found in section 2.3.2 of [4], with the result for the energy levels

$$E_{n,j} = m_e \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{1}{2}\frac{Z^4\alpha^4}{n^3}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right) + \mathcal{O}((Z\alpha)^6) \right] \tag{2.156}$$

This result predicts correctly the splitting of the energy levels with the same principle quantum number n but different j (Fine Splitting); It does not predict the observed splitting of energy levels with the same n and j but different parity $(-1)^l$ (Lamb Shift). This requires the quantization of the EM field $A^\mu$and, hence, the use of Quantum Electrodynamics (QED). Other important quantum corrections are discussed in [4].

(draw the energy levels with $n = 1$ and $n = 2$)

3 Propagator Theory

3.1 Introduction

We need a method for calculating the rates of physical processes such as scattering of particles or the decay of a particle into other particles in RQM. In non-relativistic QM this is done using time-dependent perturbation theory. We shall develope an alternative approach that can be generalized to RQM $\rightarrow$ Porpagator Theory. This is equivalent to a complete solution of the Schrödinger equation.

3.2 Non-Relativistic Propagators

The starting point is the Huygen's Principle: In Optics this says that the propagation of a light wave can be understood by assuming that each point on the wave front acts as a source of a secondary wave which spreads out spherically. This idea can be applied to QM, since the Schrödinger equation is a linear differential equation and, hence, any linear combination of known solutions is also a solution.

If the wave function $\psi(t, \vec{x})$ is known at some time t, then for some later time $t' > t$ the wave function $\psi(t', \vec{x}')$ may be found by treating each point of space at time t as a source of a spherical wave. At the time t', the amplitude of the wave which propagated from $\vec{x}$ at time t to $\vec{x}'$ at time t' will be proportional to $\psi(t, \vec{x})$. We write the contribution to $\psi(t', \vec{x}')$ as

$$iG(t', \vec{x}'; t, \vec{x})\psi(t, \vec{x}) \tag{3.1}$$

and summing over all contributions gives

$$\psi(t', \vec{x}') = i \int d^3x G(t', \vec{x}'; t, \vec{x}) \psi(t, \vec{x}) \,. \tag{3.2}$$

$G(t', \vec{x}'; t, \vec{x})$ is referred to as the Green function or Propagator, and the knowledge of $G(t', \vec{x}'; t, \vec{x})$ will enable us to construct the wave function at a later time from the wave function at an earlier time.

3.3 Construction of the Interacting Propagator from the Free Propagator

We construct the propagator for a particle that interacts with a potential an arbitrary number of times from the propagator of a freely moving particle. Later we shall construct the free propagator which we will denote by $G_0(t', \vec{x}'; t, \vec{x})$ and we will use at times x and x' as shorthand for $(t, \vec{x})$ and $(t', \vec{x}')$.

Suppose that an interaction potential is turned on at time t_1, for a short time interval Δt_1. For $t < t_1$, the wave function is the free particle wave function that we denote by $\phi(t, \vec{x})$. It obeys the equation

$$\phi(t', \vec{x}') = i \int d^3x G_0(t', \vec{x}'; t, \vec{x}) \phi(t, \vec{x}) \,. \tag{3.3}$$

For $t > t_1 + \Delta t_1$, the wave function ψ differs from ϕ by an amount $\Delta\phi$ induced by the potential V. $\Delta\phi$ can be calculated using the Schrödinger equation

$$\begin{aligned} &\text{(a)} \qquad i\frac{\partial \phi}{\partial t} = H_0 \phi \,, \\ &\text{(b)} \qquad i\frac{\partial \psi}{\partial t} = (H_0 + V)\psi \,, \end{aligned} \tag{3.4}$$

where H_0 denotes the free Hamiltonian, and (a) and (b) are the Schrödinger equation without and with interaction, respectively.

Integrating (a) and (b) from t_1 to $t_1 + \Delta t_1$, at $\vec{x} = \vec{x}_1$, yields

$$\begin{aligned} &\text{(a)} \qquad i\left(\phi(t_1 + \Delta t_1, \vec{x}_1) - \phi(t_1, \vec{x}_1)\right) = H_0 \phi(t_1, \vec{x}_1) \Delta t_1 \,, \\ &\text{(b)} \qquad i\left(\psi(t_1 + \Delta t_1, \vec{x}_1) - \psi(t_1, \vec{x}_1)\right) = H_0 \psi(t_1, \vec{x}_1) \Delta t_1 + V \psi(t_1, \vec{x}_1) \Delta t_1 \,. \end{aligned} \tag{3.5}$$

Since at time t_1 the interaction has not occured yet $\psi(t_1, \vec{x}_1) = \phi(t_1, \vec{x}_1)$, and subtracting (b)-(a) gives, provided that Δt_1 is small:

$$\begin{aligned} & i\left(\psi(t_1 + \Delta t_1, \vec{x}_1) - \phi(t_1 + \Delta t_1, \vec{x}_1)\right) = V(t_1, \vec{x}_1)\psi(t_1, \vec{x}_1)\Delta t_1 \\ \rightarrow \quad & i\Delta\phi(t_1, \vec{x}_1) = V(t_1, \vec{x}_1)\phi(t_1, \vec{x}_1)\Delta t_1 \\ \rightarrow \quad & \Delta\phi(t_1, \vec{x}_1) = -iV(t_1, \vec{x}_1)\phi(t_1, \vec{x}_1)\Delta t_1 \end{aligned} \tag{3.6}$$

For $t > t_1 + \Delta t_1$, there is no interaction and so $\Delta\phi$ evolves like a free particle wave function, so that

$$\begin{aligned} & \Delta\phi(t', \vec{x}') = i \int d^3x_1 G_0(t', \vec{x}'; t_1, \vec{x}_1) \Delta\phi(t_1, \vec{x}_1) \\ \rightarrow \quad & \Delta\phi(t', \vec{x}') = \int d^3x_1 G_0(t', \vec{x}'; t_1, \vec{x}_1) V(t_1, \vec{x}_1) \phi(t_1, \vec{x}_1) \Delta t_1 \end{aligned} \tag{3.7}$$

The complete wave funtion at $t', \vec{x}'$ for $t' > t_1 + \Delta t_1$ is therefore $\psi(t', \vec{x}') \equiv \psi(x') = \phi(t', \vec{x}') + \Delta\phi(t', \vec{x}')$ and hence

$$\begin{aligned} \psi(t', \vec{x}') \quad &= \quad i \int d^3x_1 G_0(x'; x) \phi(x) \\ & \qquad + \int d^3x_1 G_0(x'; x_1) V(x_1) \phi(x_1) \Delta t_1 \, . \end{aligned} \tag{3.8}$$

Also $\phi(x_1) = i \int d^3x G_0(x_1; x) \phi(x)$, so

$$\psi(x') = i \int d^3x \left[G_0(x'; x) + \int d^3x_1 G_0(x'; x_1) V(x_1) G_0(x_1; x) \Delta t_1 \right] \psi(x) \tag{3.9}$$

where we have used that $\phi(x) = \psi(x)$ for $t \leq t_1$. We conclude that

$$\psi(x') = i \int d^3x G(x'; x) \psi(x) \tag{3.10}$$

with

$$G(x'; x) = G_0(x'; x) + \int d^3x_1 G_0(x'; x_1) V(x_1) G_0(x_1; x) \Delta t_1 \tag{3.11}$$

This is the propagator for a single interaction between $t = t_1$ and $t = t_1 + \Delta t_1$.

Continuous Interaction: if the interaction of the particle with the potential is occuring continuously the Interaction Propagator becomes

$$G(x'; x) = G_0(x'; x) + \int d^4x_1 G_0(x'; x_1) V(x_1) G_0(x_1; x) + \ldots \tag{3.12}$$

where we have replaced Δt_1 simply by $\int dt$, where the $\ldots$ indicate higher order terms for multiple interaction that go like V^2, V^3, $\ldots$ and we have defined $\int d^4x = \int dt \int d^3x$.

The wavefunction at a $t' > t$ is related to the wave function at time t by

$$\psi(x') = i \int d^3x G(x'; x) \psi(x) \tag{3.13}$$

3.4 Scattering Amplitudes

We are interested in studying the scattering of a particle off a potential, e.g. an electron scattering off a Coulomb potential due to a nucleus.

The incoming particle for $t \to -\infty$ is a free particle, and the outgoing particle for $t' \to +\infty$ is a free particle. We want to calculate the Probability Amplitude S_{fi} for the transition from the state with the free particle wave function $\phi_i(t, \vec{x})$ for $t \to -\infty$ to the state with the free particle wave function $\phi_f(t', \vec{x}')$ for $t' \to +\infty$. If we work with plane wave solutions, then for $t \to -\infty$,

$$\phi_i(t, \vec{x}) \sim e^{-i(E_i t - \vec{p}_i \cdot \vec{x})} \tag{3.14}$$

and similarly for ϕ_f for $t' \to +\infty$.

S_{fi} is called the scattering amplitude at the time $t' > t$. The Interacting particle wave function ψ_i which develops from the free particle wave function ϕ_i is given by

$$\psi_i(x') = \lim_{t \to -\infty} i \int d^3x G(x'; x) \phi_i(x) \,. \tag{3.15}$$

If we substitute for $G(x'; x)$ from eqn. (3.12) we can calculate the Interacting particle wave function $\psi_i(x')$ to arbitrary order in V. Then,

$$\begin{aligned} S_{fi} &= \lim_{t' \to +\infty} \langle \phi_f(x') | \psi_i(x') \rangle \\ &= \lim_{t' \to +\infty} \int d^3x' \phi_f^*(x') \psi_i(x') \,. \end{aligned} \tag{3.16}$$

3.5 Differential Equation for G

To obtain G_0 (and hence G) we derive a differential equation for G which we specialize to G_0. We will then be able to calculate scattering amplitudes following section 3.3 and 3.4.

Write $\psi(x') = i \int d^3x G(x', x) \psi(x)$ for $t' > t$ and $\psi(x') = 0$ for $t' < t$ to ensure that NO propagation of waves backwards in time occurs which would violate causality. Equivalently:

$$G(x'; x) = 0 \ , \ \text{for } t' < t \,. \tag{3.17}$$

We can summarize this in the equation

$$\theta(t' - t) \psi(x') = i \int d^3x G(x'; x) \psi(x) \tag{3.18}$$

where θ is the step function, which is defined as

$$\theta(t'-t) = \begin{cases} 1\,, & t'-t>0 \\ 0\,, & t'-t<0 \end{cases} \tag{3.19}$$

The step function also has an interesting integral representation

$$\theta(\tau) = \lim_{\epsilon\to 0^+} -\frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{d\omega e^{-i\omega\tau}}{(\omega+i\epsilon)}\,. \tag{3.20}$$

This can be checked by evaluating the corresponding contour integral in the complex ω plane, where the contour is taken to be a path along the real axis which is closed at infinity by a half-circle in the lower or upper half-plane. For $\tau > 0$ integrate around a half-circle in the lower half-plane to ensure exponential damping of the integrand, and the value of the integral is 1 by Cauchy's theorem, since the pole at $\omega = -i\epsilon$ is inside the closed integration contour. For $\tau < 0$ the contour is closed above and the integral vanishes because the pole lies outside the contour.

The derivative of the step function θ is the Dirac δ function

$$\frac{d\theta}{d\tau} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{-i\omega\tau} = \delta(\tau)\,. \tag{3.21}$$

The wave function $\psi(x')$ obeys the Schrödinger equation

$$i\frac{\partial\psi(x')}{\partial t'} = H(x')\psi(x')\,, \tag{3.22}$$

thus,

$$\begin{aligned} &\left(i\frac{\partial}{\partial t'} - H(x')\right)(\theta(t'-t)\psi(x')) \\ &= i\frac{\partial\theta(t'-t)}{\partial t'}\psi(x') + \theta(t'-t)\left(i\frac{\partial}{\partial t'} - H(x')\right)\psi(x') \\ &= i\delta(t'-t)\psi(x') \\ &= i\int d^3x\left(i\frac{\partial}{\partial t'} - H(x')\right)G(x';x)\psi(x)\,. \end{aligned} \tag{3.23}$$

This is true for arbitrary ψ, thus

$$\left(i\frac{\partial}{\partial t'} - H(x')\right)G(x';x) = \delta(t'-t)\delta(\vec{x}'-\vec{x}) \equiv \delta^4(x'-x) \tag{3.24}$$

because for any function $f(\vec{x})$

$$\int d^3x\delta(\vec{x}'-\vec{x})f(\vec{x}) = f(\vec{x}')\,. \tag{3.25}$$

Eqn. (3.24) is the differential equation which we shall solve for G_0 with the boundary condition $G_0(x';x) = 0$ for $t' < t$. This defines the retarded Green's function or propagator.

3.6 Free Particle Propagator

Now we will solve (3.24) for the case of a free particle. Then,

$$H(\vec{x}) = H_0(\vec{x}) = -\frac{1}{2m}\vec{\nabla}^2\,, \tag{3.26}$$

and (3.24) becomes

$$\left(i\frac{\partial}{\partial t'} + \frac{1}{2m}\vec{\nabla'}^2\right) G_0(x';x) = \delta^4(x'-x) \tag{3.27}$$

where $\vec{\nabla}'$ denotes derivates with respect to (the components of) $\vec{x}'$. Because of the homogeneity of space-time, G_0 can only depend on the combination $x' - x$.

Perform a Fourier Transform:

$$G_0(x',x) = G_0(x'-x) = \int \frac{d\omega d^3p}{(2\pi)^4} G_0(\omega,\vec{p}) e^{-i\omega(t'-t)} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}\,. \tag{3.28}$$

Substituting this in (3.27) gives

$$\int \frac{d\omega d^3p}{(2\pi)^4}\left(\omega - \frac{\vec{p}^2}{2m}\right) G_0(\omega,\vec{p}) e^{-i\omega(t'-t)} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} = \underbrace{\int \frac{d\omega d^3p}{(2\pi)^4} e^{-i\omega(t'-t)} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}}_{=\delta^4(x'-x)}\,, \tag{3.29}$$

where we used $\frac{\partial}{\partial x'} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} = ip_x e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}$, etc.

Comparing the Fourier coefficients,

$$\begin{aligned} &\left(\omega - \frac{\vec{p}^2}{2m}\right) G_0(\omega,\vec{p}) = 1 \\ &\rightarrow G_0(\omega,\vec{p}) = \frac{1}{\left(\omega - \frac{\vec{p}^2}{2m}\right)} \end{aligned} \tag{3.30}$$

except for the singularity at $\omega = \frac{\vec{p}^2}{2m}$. The correct presription to deal with this singularity which ensures the retarded boundary condition turns out to be

$$G_0(\omega,\vec{p}) = \frac{1}{\left(\omega - \frac{\vec{p}^2}{2m} + i\epsilon\right)}\,, \tag{3.31}$$

where $\epsilon \to 0^+$ after integration.

Now,

$$G_0(x'-x) = \lim_{\epsilon\to 0^+} \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{e^{-i\omega(t'-t)}}{\left(\omega - \frac{\vec{p}^2}{2m} + i\epsilon\right)} . \tag{3.32}$$

Using the same contour integrations as for $\theta(t'-t)$ in section 3.5, we find that

$$\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{e^{-i\omega(t'-t)}}{\left(\omega - \frac{\vec{p}^2}{2m} + i\epsilon\right)} = -i\theta(t'-t) e^{-i\frac{\vec{p}^2}{2m}(t'-t)} \tag{3.33}$$

thus

$$G_0(x'-x) = -i \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} e^{-i\frac{\vec{p}^2}{2m}(t'-t)} \theta(t'-t) . \tag{3.34}$$

We can now check by inserting this expression for $G_0(x'-x)$ into (3.27) that this is indeed a solution of (3.27). Because of the $\theta(t'-t)$ factor it also satisfies the correct boundary condition $G_0(x'-x) = 0$ for $t' < t$.

3.7 Interacting Propagator and S_{fi}

We can calculate the interacting propagator from the free propagator using

$$G(x';x) = G_0(x';x) + \int d^4x_1 G_0(x';x_1) V(x_1) G_0(x_1;x) + \text{order } V^2 \tag{3.35}$$

Then we can use section 3.4 to calculate the scattering amplitude S_{fi} from the initial state i to the final state f

$$\begin{aligned} S_{fi} &= \int d^3x' \phi_f^*(x') \psi_i(x') \\ &\to S_{fi} = i \lim_{t'\to+\infty, t\to-\infty} \int d^3x' \int d^3x \phi_f^*(x') G(x';x) \phi_i(x) \end{aligned} \tag{3.36}$$

Substituting (3.35) into (3.36), the G_0 term contribution is

$$\begin{aligned} S_{fi} &= \lim_{t'\to+\infty, t\to-\infty} \int d^3x' \phi_f^*(x') i \int d^3x G_0(x';x) \phi_i(x) + \dots \\ &= \lim_{t'\to+\infty} \int d^3x' \phi_f^*(x') \phi_i(x') + \dots \\ &= \delta_{fi} + \dots , \end{aligned} \tag{3.37}$$

if we use Normalised wave functions. In the particular case of plane wave functions δ_{fi} is replaced by $\delta^3(\vec{p}_f - \vec{p}_i)$.

Thus the complete expression is

$$\begin{aligned} S_{fi} &= \delta_{fi} + i \lim_{t' \to +\infty, t \to -\infty} \int d^3x' \int d^3x \int d^4x_1 \\ &\quad \times \phi_f^*(x') G_0(x'; x_1) V(x_1) G_0(x_1; x) \phi_i(x) + \dots \\ &\Rightarrow \quad S_{fi} = \delta_{fi} - i \int d^4x_1 \phi_f^*(x_1) V(x_1) \phi_i(x_1) + \dots \end{aligned} \tag{3.38}$$

If the interaction potential V is not too large then the first few terms give a good approximation to the Scattering Amplitude.

3.8 Relativistic Electron Propagator

To generalize the propagator theory to the relativistic case we proceed somewhat intuitively. We need to generalize the differential equation for the non-relativistic propagator $G(x'; x)$, which we found in Section 3.5 to be

$$\left(i \frac{\partial}{\partial t'} - H(x') \right) G(x'; x) = \delta^4(x' - x) \,. \tag{3.39}$$

This should be compared to the usual (homogeneous) Schrödinger equation

$$\left(i \frac{\partial}{\partial t} - H(x) \right) \Psi(x) = 0 \,. \tag{3.40}$$

With the Dirac equations in an electromagnetic field

$$\begin{aligned} (i \not{\nabla} - m) \, \Psi &= -q \gamma \cdot A \Psi \\ &\Rightarrow \quad (i \not{\nabla} - e \not{A} - m) \, \Psi = 0 \,, \end{aligned} \tag{3.41}$$

this suggests that the Relativistic Propagator, which we will denote by $\widehat{S}_F(x'; x)$ should obey the differential equation

$$(i \not{\nabla}' - e \not{A}' - m) \, \widehat{S}_F(x'; x) = \mathbb{I} \delta^4(x' - x) \,, \tag{3.42}$$

where $\widehat{S}_F$ is a 4×4 matrix. In longhand notation this becomes

$$\sum_{\lambda=1}^{4} \left[\gamma^\mu (i \nabla'_\mu - e A_\mu(x')) - m \mathbb{I} \right]_{\alpha\lambda} \left[\widehat{S}_F(x'; x) \right]_{\lambda\beta} = \mathbb{I}_{\alpha\beta} \delta^4(x' - x) \,. \tag{3.43}$$

3.9 Free Electron Propagator

The free electron Propagator is denoted by $S_F(x';x)$ and obeys the differential equation

$$(i\not{\nabla}'- m)\, S_F(x';x) = \mathbb{I}\delta^4(x'-x)\,. \tag{3.44}$$

Using homogeneity of space-time we can take S_F to depend only on $x'-x$. Now write S_F as a Fourier transform

$$S_F(x';x) = S_F(x'-x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)} \tilde{S}_F(p) \tag{3.45}$$

and substitute it back into (3.44). Noticing that $p\cdot(x'-x) = p_0(t'-t) - \vec{p}\cdot(\vec{x}'-\vec{x})$ and $\not{\nabla}'= \gamma^0\frac{\partial}{\partial t'} + \vec{\gamma}\cdot\vec{\nabla}'$ we get

$$\int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)}(\gamma^0 p_0 - \vec{\gamma}\cdot\vec{p} - m)\tilde{S}_F(p) = \delta^4(x'-x)\mathbb{I}\,. \tag{3.46}$$

But on the other hand the Fourier transform of a Dirac δ function is 1, so we can write

$$\delta^4(x'-x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)}\,, \tag{3.47}$$

and we only have to compare Fourier coefficients. We conclude that

$$(\not{p} - m)\tilde{S}_F(p) = \mathbb{I}\,, \tag{3.48}$$

and hence we find for the free electron propagator in momentum space

$$\tilde{S}_F(p) = (\not{p} - m)^{-1}\,, \tag{3.49}$$

which can also be written as

$$\tilde{S}_F(p) = \frac{(\not{p} + m)}{(p^2 - m^2)}\,, \tag{3.50}$$

for $p^2 = m^2$.

There are singularities at $p^2 = m^2$ because the denominator in $\tilde{S}_F(p)$ becomes zero and the propagator develops poles. The singularities are localized at

$$\begin{aligned} p_0^2 - \vec{p}^2 &= m^2 \\ \Rightarrow\ p_0 &= \pm\sqrt{\vec{p}^2 + m^2} \equiv \pm E\,. \end{aligned} \tag{3.51}$$

This means that the singularities occur when the 4-momentum p is *on-shell* i.e. it obeys the relativistic energy-momentum relation $p^2 = m^2$.

We use Feynman's boundary conditions to handle the singularities. More specifically, $S_F(x'-x)$ describes the the propagation of an electron (e^-) from the space-time point

x to x'. Positive (negative) energy electrons are represented by wave functions with positive (negative) frequency time behaviour, namely $e^{-i\omega t}$ ($e^{i\omega t}$) where $\omega > 0$. Because of hole theory, electrons are associated with positive energy electrons propagating *forward in time*, and positrons are associated with negative energy wave functions propagating backwards in time. (Recall that a negative energy wave function of 4-momentum $-p$, which is propagating backwards in time, represents a positron of momentum $+p$, which is therefore propagating *forward in time*.)

To ensure that an electron (positron) does not spontaneously change into a positron (electron) we require that a positive (negative) frequency wave propagating from x into the future (past) possesses ONLY positive (negative) frequency components. Therefore we have to require that for $t' > t$ ($t' < t$) the propagator $S_F(x' - x)$ contains only positive (negative) frequency components. These are the Feynman boundary conditions on S_F; it turns out that in momentum space this translates into ($i\epsilon$ prescription)

$$\tilde{S}_F(p) = \frac{(\not{p} + m)}{(p^2 - m^2 + i\epsilon)}\,, \tag{3.52}$$

where $\epsilon \to 0^+$.

Let us check this for $t' > t$. Starting point is the Fourier transform of (3.52)

$$\begin{aligned} S_F(x' - x) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)} \frac{(\not{p} + m)}{(p^2 - m^2 + i\epsilon)} \\ &= \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} \int \frac{dp_0}{2\pi} e^{-ip_0(t'-t)} \frac{(\not{p} + m)}{(p^2 - m^2 + i\epsilon)}\,. \end{aligned} \tag{3.53}$$

Let us focus on the p_0 integration. For $t' > t$ we evaluate the p_0 integral by closing the contour with a semi-circle in the lower half p_0-plane, since

$$\begin{aligned} e^{-ip_0(t'-t)} &= e^{-i\Re(p_0)(t'-t)+\Im(p_0)(t'-t)} \\ &= e^{\Im(p_0)(t'-t)} e^{-i\Re(p_0)(t'-t)} \end{aligned} \tag{3.54}$$

and for $t' > t$ we have $e^{\Im(p_0)(t'-t)} \to 0$ if $\Im(p_0) \to -\infty$.

Thus the integral along the large semi-circle is zero. Therefore,

$$\begin{aligned} &\int_{-\infty}^{+\infty} \frac{dp_0}{2\pi} e^{-ip_0(t'-t)} \frac{(\not{p} + m)}{(p^2 - m^2 + i\epsilon)} \\ = \; &-\int_{\mathcal{C}} \frac{dp_0}{2\pi} e^{-ip_0(t'-t)} \frac{(\not{p} + m)}{(p^2 - m^2)} \end{aligned} \tag{3.55}$$

where $\mathcal{C}$ denotes the closed contour. For small ϵ, the singularities are at $p_0 = E - \frac{i\epsilon}{2E}$ and $p_0 = -E + \frac{i\epsilon}{2E}$. Since $\epsilon > 0$, the singularity at $p_0 = -E$ is pushed into the upper half p_0 plane and is, therefore, outside the contour and the singularity at $p_0 = E$ is inside the

contour. Instead of adding the small imaginary parts, we can slightly deform the contour to include the singularity at $p = E$ and exclude the singularity at $p_0 = -E$.

$$e^{-ip_0(t'-t)}\frac{(\not{p}+m)}{(p^2-m^2)} = \frac{(p_0\gamma^0 - \vec{p}\cdot\vec{\gamma}+m)e^{-ip_0(t'-t)}}{(p_0+E)(p_0-E)} \tag{3.56}$$

with $E = \sqrt{\vec{p}^2+m^2}$.

Thus the residue at $p_0 = E$ is

$$\frac{(E\gamma^0 - \vec{p}\cdot\vec{\gamma}+m)e^{-iE(t'-t)}}{2E}\,. \tag{3.57}$$

Consequently, for $t' > t$

$$S_F(x'-x) = -i\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}e^{-iE(t'-t)}\frac{(E\gamma^0-\vec{p}\cdot\vec{\gamma}+m)}{2E} \tag{3.58}$$

As required this contains only positive frequency modes.

For $t' < t$ we have to close the contour in the upper half p_0 plane and we find in a similar fashion that only negative frequency components contribute.

3.10 Electron Propagator in an Electromagnetic Field

The propagator $\widehat{S}_F(x';x)$ satisfies

$$(i\not{\nabla}'- e\not{A}'-m)\widehat{S}_F(x';x) = \delta^4(x'-x)\mathbb{I} \tag{3.59}$$

We obtain an integral equation for $\widehat{S}_F(x';x)$ as follows

$$\begin{aligned}&(i\not{\nabla}'- m)\widehat{S}_F(x';x) = e\not{A}'\widehat{S}_F(x';x) + \delta^4(x'-x)\\ \Rightarrow\quad &(i\not{\nabla}'- m)\widehat{S}_F(x';x) = \int d^4x''\delta^4(x''-x')\left[\delta^4(x''-x) + e\not{A}''\widehat{S}_F(x'';x)\right]\end{aligned} \tag{3.60}$$

However, $(i\not{\nabla}'- m)S_F(x';x) = \delta^4(x'-x)$ for the free propagator. Consistency is obtained if

$$\widehat{S}_F(x';x) = \int d^4x''S_F(x';x'')\left[\delta^4(x''-x) + e\not{A}''\widehat{S}_F(x'';x)\right]\,, \tag{3.61}$$

which can be checked by acting on both sides with $(i\not{\nabla}'- m)$.

Therefore, replacing x'' by x_1, we obtain

$$\widehat{S}_F(x';x) = S_F(x';x) + e\int d^4x_1 S_F(x';x_1)\not{A}(x_1)\widehat{S}_F(x_1;x)\,, \tag{3.62}$$

which is an integral equation for $\widehat{S}_F$ which may be solved by iteration;

- zeroth order: Neglecting $e\not{A}$ we have $\widehat{S}_F(x';x) = S_F(x';x)$.
- first order: To first order in $e\not{A}$, substitute the zeroth order solution for $\widehat{S}_F$ on the righthand side of (3.62) to obtain

$$\widehat{S}_F(x';x) = S_F(x';x) + e\int d^4x_1 S_F(x';x_1)\not{A}(x_1)S_F(x_1;x)\,, \tag{3.63}$$

  and so on for higher orders.

The positive energy Dirac spinor $\widehat{\Psi}$ at time $t' > t$ is related to the wave function Ψ at time t by

$$\widehat{\Psi}(x') = i\int d^3x \widehat{S}_F(x';x)\Psi(x)\,, \tag{3.64}$$

where $\widehat{\Psi}$ denotes an *Interacting electron*, whereas the free wave function at $t' > t$ for a positive energy electron is

$$\Psi(x') = i\int d^3x S_F(x';x)\Psi(x)\,. \tag{3.65}$$

3.11 Scattering Amplitudes for Electrons

The incoming particle for $t \to -\infty$ is a free, positive energy electron and the outgoing particle for $t' \to +\infty$ is a free electron. The electron has been scattered by the electromagnetic field A_μ. We wish to calculate the probability amplitude S_{fi} for the transition from the state with the free electron wave function $\psi_i(t,\vec{x})$ for $t \to -\infty$ to the state with the electron wavefunction $\psi_f(t',\vec{x}')$ for $t' \to +\infty$.

At time $t' > t$, the interacting electron wavefunction $\widehat{\psi}_i$ which develops from the free electron wave function ψ_i is given by

$$\widehat{\psi}_i(x') = \lim_{t\to-\infty} i\int d^3x \widehat{S}_F(x';x)\psi_i(x)\,. \tag{3.66}$$

Substituting for $\widehat{S}_F$ from (3.62) we get $\widehat{\psi}_i(x')$ to any required order in e. Then

$$\begin{aligned} S_{fi} &= \lim_{t'\to+\infty}\langle\psi_f(x')|\widehat{\psi}_i(x')\rangle \\ &= \lim_{t'\to+\infty}\int d^3x'\overline{\psi}_f(x')\widehat{\psi}_i(x')\,, \end{aligned} \tag{3.67}$$

where $\overline{\psi}$ denotes the adjoint spinor $\psi^\dagger\gamma^0$.

Proceeding as before

$$S_{fi} = i \lim_{t' \to +\infty, t \to -\infty} \int d^3x' \int d^3x \overline{\psi}_f(x') \widehat{S}_F(x'; x) \psi_i(x) \tag{3.68}$$

Substituting (3.63) for $\widehat{S}_F$ in (3.68) gives

$$\begin{aligned} S_{fi} &= \lim_{t' \to +\infty, t \to -\infty} \int d^3x' \overline{\psi}_f(x') i \int d^3x S_F(x'; x) \psi_i(x) + \ldots \\ &= \lim_{t' \to +\infty} \int d^3x' \overline{\psi}_f(x') \psi_i(x') + \ldots = \delta_{fi} + \ldots \end{aligned} \tag{3.69}$$

Note, that δ_{fi} has to be replaced by a Dirac δ function if we consider plane wave solutions.

If we use normalized wave functions the complete expression (to first order in the interaction) for S_{fi} becomes

$$\begin{aligned} S_{fi} &= \delta_{fi} + i \lim_{t' \to +\infty, t \to -\infty} e \int d^3x' \int d^3x \int d^4x_1 \overline{\psi}_f(x') S_F(x'; x_1) \not{A}(x_1) S_F(x_1; x) \psi_i(x) + \ldots \\ &\Rightarrow \quad S_{fi} = \delta_{fi} - ie \int d^4x_1 \overline{\psi}_f(x_1) \not{A}(x_1) \psi_i(x_1) \end{aligned} \tag{3.70}$$

3.12 Scattering Amplitudes Involving Positrons

For a positron scattering off an electromagnetic field, we note that to lowest order in the interaction the cross section is identical to electron scattering. Here the incoming state is a negative energy electron in the future propagating backwards in time with wave function ψ_i and the outgoing state is a negative energy electron in the past with wave function ψ_f. Then,

$$S_{fi} = \lim_{t' \to -\infty} \langle \psi_f(x') | \widehat{\psi}_i(x') \rangle \tag{3.71}$$

and we get essentially the same result with negative energy wavefunctions. Note that the momentum of the "incoming" negative energy eletron is related to the momentum of the corresponding outgoing positron via $p_i^{e^-} = -p_f^{e^+}$, and that the momentum of the "outgoing" negative energy eletron is related to the momentum of the corresponding outgoing positron via $p_f^{e^-} = -p_i^{e^+}$.

4 Applications

4.1 Scattering of an Electron from a Fixed Coulomb Potential

This is the Rutherford scattering problem. An electron e^- scatters off the Coulomb potential due to a nucleus. We calculate the scattering amplitude S_{fi} to first order in e

(i.e. to first order in the interaction) then, if the final state is different from the initial state ($i \neq f$)

$$S_{fi} = -ie \int d^4x \overline{\psi}_f(x) \not{A}(x) \psi_i(x) \tag{4.1}$$

If we choose free particle wave functions ψ_i and ψ_f normalised in a box of volume V, then

$$\begin{aligned} \psi_i(x) &= \frac{1}{\sqrt{2E_iV}} e^{-ip_i\cdot x} U(p_i, s_i) \text{ and} \\ \psi_f(x) &= \frac{1}{\sqrt{2E_fV}} e^{-ip_f\cdot x} U(p_f, s_f)\,. \end{aligned} \tag{4.2}$$

Normalisation of wave functions

Recall that the probability density is $\rho = \psi^\dagger\psi$. Now take

$$\begin{aligned} \psi &= N e^{-ip\cdot x} U(p,s) \\ \Rightarrow\ \rho &= \psi^\dagger\psi = |N|^2 e^{-ip\cdot x} e^{+ip\cdot x} U^\dagger(p,s)U(p,s) = |N|^2 U^\dagger(p,s)U(p,s) \end{aligned} \tag{4.3}$$

Integrate this over the volume V

$$\begin{aligned} \int d^3x \psi^\dagger\psi &= |N|^2 U^\dagger(p,s)U(p,s) \int d^3x \\ &= |N|^2 U^\dagger(p,s)U(p,s)V \\ &= |N|^2 2EV \equiv 1\,, \end{aligned} \tag{4.4}$$

from which we find

$$|N|^2 = \frac{1}{2EV} \Rightarrow N = \frac{1}{\sqrt{2EV}}\,. \tag{4.5}$$

To describe the Coulomb potential due to a nucleus of atomic number Z we take

$$A_0 = \frac{-Ze}{4\pi|\vec{x}|}\ ,\ \vec{A} = 0\,. \tag{4.6}$$

Then

$$S_{fi} = \frac{ie^2Z}{4\pi}\frac{1}{2V}\frac{1}{\sqrt{E_iE_f}} \overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i) \underbrace{\int \frac{d^4x}{|\vec{x}|} e^{i(p_f - p_i)\cdot x}}_{=I}\,, \tag{4.7}$$

and the integral is

$$\begin{aligned} I &= \int dx^0 e^{i(E_f - E_i)x^0} \int \frac{d^3x}{|\vec{x}|} e^{-i(\vec{p}_f - \vec{p}_i)\cdot\vec{x}} \\ &= 2\pi\delta(E_f - E_i)\frac{4\pi}{|\vec{q}|^2} \end{aligned} \tag{4.8}$$

where $\vec{q} = \vec{p}_f - \vec{p}_i$.

Thus the scattering amplitude is

$$S_{fi} = \frac{ie^2 Z}{2V\sqrt{E_i E_f}} \frac{\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)}{|\vec{q}^2|} 2\pi\delta(E_f - E_i)\,. \tag{4.9}$$

The transition probability from state i to state f is given by $|S_{fi}|^2$

$$|S_{fi}|^2 = \frac{Z^2(4\pi\alpha)^2}{4V^2 E_i E_f} \frac{|\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2}{|\vec{q}|^4} [2\pi\delta(E_f - E_i)]^2 \tag{4.10}$$

What is the interpretation of $[2\pi\delta(E_f - E_i)]^2$?

Write

$$\begin{aligned} 2\pi\delta(E_f - E_i) &= \int_{-\infty}^{+\infty} dt e^{i(E_f - E_i)t} \\ &= \lim_{T\to\infty} \int_{-T/2}^{+T/2} dt e^{i(E_f - E_i)t} \\ &= \lim_{T\to\infty} \frac{1}{i(E_f - E_i)} \left[e^{i(E_f - E_i)t}\right]_{-T/2}^{T/2} \\ &= \lim_{T\to\infty} \frac{2\sin((E_f - E_i)T/2)}{(E_f - E_i)}\,, \end{aligned} \tag{4.11}$$

thus

$$[2\pi\delta(E_f - E_i)]^2 = \lim_{T\to\infty} \frac{4\sin^2((E_f - E_i)T/2)}{(E_f - E_i)^2}\,. \tag{4.12}$$

However, it is known that for $\alpha < E_i < \beta$

$$\int_\alpha^\beta \frac{\sin^2((E_f - E_i)T/2)}{(E_f - E_i)^2} dE_f = 2\pi T\,, \tag{4.13}$$

for large T. Hence, we can interprete the RHS of (4.12) as $2\pi T\delta(E_f - E_i)$ and we can take the intepretation

$$[2\pi\delta(E_f - E_i)]^2 = 2\pi T\delta(E_f - E_i)\,. \tag{4.14}$$

The difficulty we encountered can be avoided by using wave packets instead of plane waves.

With this intepretation of the square of the Dirac δ-function, the transition probability per unit time becomes

$$\frac{|S_{fi}|^2}{T} = \frac{8\pi^3 Z^2\alpha^2}{V^2 E_i E_f} \frac{|\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2}{|\vec{q}|^4} \delta(E_f - E_i)\,. \tag{4.15}$$

We are interested in the transition probability with a chosen range of final momenta, $\vec{p}_f$ to $\vec{p}_f + d\vec{p}_f$ containing $\frac{V}{(2\pi)^3} d^3p_f$ states. This is

$$\frac{V}{(2\pi)^3} d^3p_f |S_{fi}|^2 \,. \tag{4.16}$$

In spherical coordinates in momentum space

$$d^3p_f = p_f^2 dp_f \underbrace{\sin\theta d\theta d\phi}_{d\Omega = \text{ solid angle}} \,, \tag{4.17}$$

where $p_f = |\vec{p}_f|$. Thus, we require

$$\frac{V}{(2\pi)^3} p_f^2 dp_f d\Omega \frac{|S_{fi}|^2}{T} \tag{4.18}$$

The differential cross section $d\sigma(\theta, \phi)$ for scattering into the final solid angle $d\Omega$ is defined to be

$$d\sigma(\theta,\phi) = \frac{\textit{Probability per unit time of a transition into } d\Omega}{\textit{Flux of incident particles}} \tag{4.19}$$

where *(Flux of incident particles)* = (Incident probability per unit area per unit time).

Thus,

$$\text{Flux} = \underbrace{(\psi_i^\dagger \psi_i)}_{\textit{Probability density}} \times \underbrace{\frac{|\vec{p}_i|}{E_i}}_{=|\vec{v}_i|} \tag{4.20}$$

Recall probability density $\rho = \psi^\dagger \psi$ and with our normalisation

$$\psi_i^\dagger \psi_i = 2E|N|^2 = \frac{2E}{2EV} = \frac{1}{V} \,. \tag{4.21}$$

Thus, the flux of incident particles is

$$\text{Flux} = \frac{|\vec{p}_i|}{VE_i} \,. \tag{4.22}$$

Now the differential cross section into a solid angle $d\Omega$ is given by

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{8\pi^3 V Z^2 \alpha^2 E_i}{(2\pi)^3 V E_i |\vec{p}_i|} \int_0^\infty \frac{|\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2}{|\vec{q}|^4} \delta(E_f - E_i) \frac{p_f^2 dp_f}{E_f} \\ &= \frac{Z^2\alpha^2}{|\vec{p}_i|} \int_0^\infty \frac{|\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2}{|\vec{q}|^4} \delta(E_f - E_i) \frac{p_f^2 dp_f}{E_f} \,. \end{aligned} \tag{4.23}$$

Because $E_f^2 = p_f^2 + m^2$ we have

$$2E_f dE_f = 2p_f dp_f \,, \tag{4.24}$$

and the integral in (4.23) becomes

$$\int_0^\infty dE_f \frac{|\overline{U}\gamma^0 U|^2}{|\vec{q}|^4} p_f \delta(E_f - E_i) \,. \tag{4.25}$$

The δ function results in setting $E_f = E_i$ and $|\vec{p}_f| = |\vec{p}_i|$. Therefore,

$$\frac{d\sigma}{d\Omega} = \frac{Z^2\alpha^2}{|\vec{q}|^4} |\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2 \,, \tag{4.26}$$

all evaluated at $E_f = E_i$ and $|\vec{p}_f| = |\vec{p}_i|$.

If we do not observe the spin state of the final electrons, then we should sum over final spin states (polarisations) in $\frac{d\sigma}{d\Omega}$. If in the incident beam of electrons each spin is equally likely (unpolarised beam), then we should also average over the initial spin states. In that case,

$$\frac{d\sigma}{d\Omega} = \frac{Z^2\alpha^2}{2|\vec{q}|^4} \sum_{s_f, s_i} |\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)|^2 \,, \tag{4.27}$$

all evaluated at $E_f = E_i$ and $|\vec{p}_f| = |\vec{p}_i|$. Note that we multiplied with $1/2$ for the average over the two initial spin states.

Such spin summations are evaluated by reducing the problem to evaluating the trace of a matrix and occur often in scattering problems, so we will digress a little into this subject.

4.2 Trace Theorems

Abbreviating $\overline{U}(p_f, s_f)\gamma^0 U(p_i, s_i)$ to $\overline{U}(f)\gamma^0 U(i)$ we need to study $|\overline{U}(f)\gamma^0 U(i)|^2$. Consider the general case, which is useful for other scattering amplitudes, $|\overline{U}(f)\Gamma U(i)|^2$ where Γ is some 4×4 matrix. Usually Γ is a product of various γ-matrices.

First notice that

$$(\overline{U}(f)\Gamma U(i))^* = \overline{U}(i)\overline{\Gamma}U(f) \tag{4.28}$$

where $\overline{\Gamma} = \gamma^0\Gamma^\dagger\gamma^0$. Thus, we need to evaluate

$$|\overline{U}(f)\Gamma U(i)|^2 = \overline{U}(f)\Gamma U(i)(\overline{U}(f)\Gamma U(i))^* = \overline{U}(f)\Gamma U(i)\overline{U}(i)\overline{\Gamma}U(f) \,. \tag{4.29}$$

Using the properties of γ-matrices we can show that

$$\begin{aligned} \overline{\gamma}^{\mu} &= \gamma^{\mu} \\ \overline{\gamma}^{5} &= -\gamma^{5} \\ \overline{\gamma^{\mu}\gamma^{5}} &= \gamma^{\mu}\gamma^{5} = -\gamma^{5}\gamma^{\mu} \end{aligned} \tag{4.30}$$

and

$$\overline{\not{a}\not{b}\not{c}\ldots\not{q}} = \not{q}\ldots\not{c}\not{b}\not{a}\,. \tag{4.31}$$

The spin sums are reduced to matrix traces as follows. We use the following property of Dirac spinors which can be derived from their explicit representation

$$\sum_{s} U_{\alpha}(p,s)\overline{U}_{\beta}(p,s) = (\not{p} + m\mathbb{I})_{\alpha\beta}\,. \tag{4.32}$$

$$\begin{aligned} \textstyle\sum_{s_f,s_i} |\overline{U}(f)\Gamma U(i)|^2 &= \sum_{s_f,s_i} \overline{U}(f)\Gamma U(i)\overline{U}(i)\overline{\Gamma}U(f) \\ &= \sum_{s_f,s_i} \overline{U}(f)_{\alpha}\Gamma_{\alpha\beta}U(i)_{\beta}\overline{U}(i)_{\gamma}\overline{\Gamma}_{\gamma\delta}U(f)_{\delta} \\ &= \sum_{s_f,s_i} U(f)_{\delta}\overline{U}(f)_{\alpha}\Gamma_{\alpha\beta}U(i)_{\beta}\overline{U}(i)_{\gamma}\overline{\Gamma}_{\gamma\delta} \\ &= (\not{p}_f + m\mathbb{I})_{\delta\alpha}\Gamma_{\alpha\beta}(\not{p}_i + m\mathbb{I})_{\beta\gamma}\overline{\Gamma}_{\gamma\delta} \\ &= \sum_{\delta} \left((\not{p}_f + m\mathbb{I})\Gamma(\not{p}_i + m\mathbb{I})\overline{\Gamma}\right)_{\delta\delta} \end{aligned} \tag{4.33}$$

$$\Rightarrow \sum_{s_f,s_i} |\overline{U}(f)\Gamma U(i)|^2 = Tr\left((\not{p}_f + m\mathbb{I})\Gamma(\not{p}_i + m\mathbb{I})\overline{\Gamma}\right) \tag{4.34}$$

Therefore, the problem of evaluating the spin sums reduces to evaluating the trace of a matrix.

The following Theorems turn out to be useful in applications.

Theorem 1: $Tr(\not{a}_1\ldots\not{a}_n) = 0$ if n is odd.

Proof:

$$\begin{aligned} Tr(\not{a}_1\ldots\not{a}_n) &= Tr(\not{a}_1\ldots\not{a}_n\gamma_5\gamma_5) \\ &= Tr(\gamma_5\not{a}_1\ldots\not{a}_n\gamma_5) \\ &= (-1)^n Tr(\not{a}_1\ldots\not{a}_n\gamma_5\gamma_5) \\ &= (-1)^n Tr(\not{a}_1\ldots\not{a}_n)\,, \end{aligned} \tag{4.35}$$

where we have used $\gamma_5\gamma_5 = \mathbb{I}$ in the first line, the cyclic property of any matrix trace in the second line. The identity $\gamma_5\gamma^{\mu} = -\gamma^{\mu}\gamma_5$ was used n times in the third line to bring

γ_5 from the left in the matrix trace to the right inside the matrix trace and in line four we used again $\gamma_5\gamma_5 = \mathbb{I}$.

Thus for odd n

$$Tr(\not{a}_1 \ldots \not{a}_n) = -Tr(\not{a}_1 \ldots \not{a}_n) \Rightarrow Tr(\not{a}_1 \ldots \not{a}_n) = 0\,. \tag{4.36}$$

Theorem 2: $Tr(\not{a}\not{b}) = 4a \cdot b$

Proof:

$$\begin{aligned} Tr(\not{a}\not{b}) &= Tr(\not{b}\not{a}) \\ &= \frac{1}{2}(Tr(\not{a}\not{b}) + Tr(\not{b}\not{a})) \\ &= \frac{1}{2}Tr(\not{a}\not{b} + \not{b}\not{a}) \\ &= \frac{1}{2}Tr(a_\mu b_\nu(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)) \\ &= \frac{1}{2}Tr(a_\mu b_\nu 2g^{\mu\nu}\mathbb{I}) \\ &= a \cdot b\, Tr\mathbb{I} \\ &= 4a \cdot b \quad Q.E.D. \end{aligned} \tag{4.37}$$

Theorem 3: $Tr(\not{a}\not{b}\not{c}\not{d}) = 4[(a \cdot b)(c \cdot d) + (a \cdot d)(b \cdot c) - (a \cdot c)(b \cdot d)]$

Proof:

Use $\{\not{a}_1, \not{a}_2\} = 2a_1 \cdot a_2\mathbb{I}$. This follows from $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{I}$:

$$\begin{aligned} \{\not{a}_1, \not{a}_2\} &= \not{a}_1\not{a}_2 + \not{a}_2\not{a}_1 \\ &= (a_1)_\mu(a_2)_\nu(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) \\ &= (a_1)_\mu(a_2)_\nu 2g^{\mu\nu}\mathbb{I} \\ &= 2a_1 \cdot a_2\mathbb{I} \end{aligned} \tag{4.38}$$

$$\begin{aligned} Tr(\not{a}\not{b}\not{c}\not{d}) &= -Tr(\not{b}\not{a}\not{c}\not{d}) + 2(a \cdot b)Tr(\mathbb{I}\not{c}\not{d}) \\ &= Tr(\not{b}\not{c}\not{a}\not{d}) - 2(a \cdot c)Tr(\not{b}\mathbb{I}\not{d}) + 2(a \cdot b)Tr(\not{c}\not{d}) \\ &= -Tr(\not{b}\not{c}\not{d}\not{a}) + 2(a \cdot d)Tr(\not{b}\not{c}\mathbb{I}) - 2(a \cdot c)Tr(\not{b}\not{d}) + 2(a \cdot b)Tr(\not{c}\not{d})\,, \end{aligned} \tag{4.39}$$

from which follows

$$\begin{aligned} Tr(\not{a}\not{b}\not{c}\not{d}) + Tr(\not{b}\not{c}\not{d}\not{a}) &= 2Tr(\not{a}\not{b}\not{c}\not{d}) \\ &= 2(a \cdot d)Tr(\not{b}\not{c}) - 2(a \cdot c)Tr(\not{b}\not{d}) + 2(a \cdot b)Tr(\not{c}\not{d})\,. \end{aligned} \tag{4.40}$$

Using Theorem 2 we find

$$Tr(\not{a}\not{b}\not{c}\not{d}) = 4(a\cdot b)(c\cdot d) + 4(a\cdot d)(b\cdot c) - 4(a\cdot c)(b\cdot d)\,. \tag{4.41}$$

4.3 The Mott Cross-Section

Return to the spin avaraged differential cross-section for scattering an electron off a fixed Coulomb potential — the Mott Cross-Section.

We need $\overline{\gamma^0} = \gamma^0(\gamma^0)^\dagger\gamma^0 = \gamma^0\gamma^0\gamma^0 = \gamma^0$. Then in $d\sigma/d\Omega$ we have the term

$$\begin{aligned}
\sum_{s_i,s_f} |\overline{U}(p_f,s_f)\gamma^0 U(p_i,s_i)|^2 &= Tr\left((\not{p}_f + m\mathbb{I})\gamma^0(\not{p}_i + m\mathbb{I})\overline{\gamma^0}\right) \\
&= \left((\not{p}_f + m\mathbb{I})\gamma^0(\not{p}_i + m\mathbb{I})\gamma^0\right) \\
&= m^2 Tr(\gamma^0\gamma^0) + mTr(\not{p}_f\gamma^0\gamma^0) + mTr(\gamma^0\not{p}_i\gamma^0) + Tr(\not{p}_f\gamma^0\not{p}_i\gamma^0) \\
&= m^2 Tr(\mathbb{I}) + Tr(\not{p}_f\gamma^0\not{p}_i\gamma^0) \\
&= 4m^2 + Tr(\not{p}_f\gamma^0\not{p}_i\gamma^0)\,,
\end{aligned} \tag{4.42}$$

then introduce the 4-vector $e \equiv (1,0,0,0)$, so that $\not{e} = \gamma^0$. Hence, we can write

$$\begin{aligned}
Tr(\not{p}_f\gamma^0\not{p}_i\gamma^0) &= Tr(\not{p}_f\not{e}\not{p}_i\not{e}) \\
&= 4\left[(p_f\cdot e)(p_i\cdot e) + (p_i\cdot e)(p_f\cdot e) - (p_f\cdot p_i)(e\cdot e)\right] \\
&= 4[2E_fE_i - p_f\cdot p_i]\,.
\end{aligned} \tag{4.43}$$

Therefore,

$$\sum_{s_i,s_f} |\overline{U}(p_f,s_f)\gamma^0 U(p_i,s_i)|^2 = 4[m^2 + 2E_fE_i - p_f\cdot p_i]\,, \tag{4.44}$$

evaluated at $E_f = E_i$ and $|\vec{p}_f| = |\vec{p}_i|$ becomes,

$$\begin{aligned}
\sum_{s_i,s_f} |\overline{U}(p_f,s_f)\gamma^0 U(p_i,s_i)|^2 &= 4[m^2 + 2E_fE_i - (E_fE_i - \vec{p}_f\cdot\vec{p}_i)] \\
&= 4[m^2 + E_fE_i + \vec{p}_f\cdot\vec{p}_i] \\
&= 4[m^2 + E_fE_i + |\vec{p}_f||\vec{p}_i|\cos\theta] \\
&= 4[\underbrace{m^2}_{=E^2-|\vec{p}|^2} + E^2 + |\vec{p}|^2\cos\theta] \\
&= 4[2E^2 - |\vec{p}|^2\underbrace{(1-\cos\theta)}_{=2\sin^2(\theta/2)}] \\
&= 8[E^2 - |\vec{p}|^2\sin^2(\theta/2)]
\end{aligned} \tag{4.45}$$

$$\Rightarrow \sum_{s_i,s_f} |\overline{U}(p_f,s_f)\gamma^0 U(p_i,s_i)|^2 = 8E^2[1-\beta^2\sin^2(\theta/2)]\,, \tag{4.46}$$

with $\beta = \frac{|\vec{p}|}{E}$.

The differential cross section also depends on

$$\begin{aligned} \vec{q}^2 = (\vec{p}_f - \vec{p}_i)^2 &= \vec{p}_f^2 + \vec{p}_i^2 - 2\vec{p}_f\cdot\vec{p}_i \\ &= |\vec{p}_f|^2 + |\vec{p}_i|^2 - 2|\vec{p}_f||\vec{p}_i|\cos\theta \\ &= 2|\vec{p}|^2(1-\cos\theta) \\ &= 4|\vec{p}|^2\sin^2(\theta/2)\,, \end{aligned} \tag{4.47}$$

hence,

$$|\vec{q}|^4 = 16|\vec{p}|^4\sin^4(\theta/2)\,. \tag{4.48}$$

Now, the spin avaraged differential cross-section is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{Z^2\alpha^2}{2|\vec{q}|^4}\sum_{s_i,s_f}|\overline{U}(p_f,s_f)\gamma^0U(p_i,s_i)|^2 \\ &= \frac{Z^2\alpha^2}{32|\vec{p}|^4\sin^4(\theta/2)}8E^2[1-\beta^2\sin^2(\theta/2)] \\ &= \frac{Z^2\alpha^2}{4|\vec{p}|^2\beta^2\sin^4(\theta/2)}(1-\beta^2\sin^2(\theta/2))\,, \end{aligned} \tag{4.49}$$

where $\beta = \frac{|\vec{p}|}{E}$. This is the Mott Cross-Section.

In the non-relativistic limit, $\vec{p}\to m\vec{v}$, $\beta = \frac{|\vec{p}|}{E} = \frac{m|\vec{v}|}{m} = |\vec{v}|$ (note that we set $c=1$) and the $\beta^2\sin^2(\theta/2)$ term can be neglected. Then, we obtain the Rutherford formula for Coulomb scattering of an electron:

$$\frac{d\sigma}{d\Omega} = \frac{Z^2\alpha^2}{4m^2|\vec{v}|^4\sin^4(\theta/2)} \tag{4.50}$$

4.4 Electron Scattering from a Dirac Proton

Instead of considering the scattering in a fixed potential, as in the last section, we consider the scattering of an electron in the EM field generated by a free Dirac proton. We visualise this process in terms of the electron e^- and the proton p exchanging a photon which may be virtual, i.e. $q^2 \neq 0$

(figure goes here)

The Scattering Amplitude to first order in the charge e is given by

$$S_{fi} = -ie \int d^4x \overline{\psi}_f(x) \not{A}(x) \psi_i(x) \tag{4.51}$$

where ψ_i and ψ_f are free particle wave functions for the initial and final electron and the EM 4-vector potential A^μ is produced by the EM 4-vector current J^μ due to the proton p. Note that J^μ has components $(\rho, \vec{j})$ where ρ is the charge density and $\vec{j}$ is the current density vector. We need to calculate A^μ in terms of J^μ, which is done using the a photon propagator (Green's function).

The Maxwell equation in Lorentz gauge $\partial_\mu A^\mu = 0$ is given by

$$\Box A^\mu(x) \equiv \partial_\nu \partial^\nu A^\mu(x) = J^\mu(x)\,. \tag{4.52}$$

In the absence of any sources or currents

$$\partial_\nu \partial^\nu A^\mu(x) = 0\,, \tag{4.53}$$

which suggests the introduction of a propagator $D_F(x-y)$ for the photon, obeying

$$\partial_\nu \partial^\nu D_F(x-y) = \delta^4(x-y)\,, \tag{4.54}$$

where the derivatives are taken w.r.t. x. The corresponding momentum space propagator $\tilde{D}_F(q)$ is defined by

$$D_F(x-y) = \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} \tilde{D}_F(q)\,. \tag{4.55}$$

Therefore,

$$\partial_\nu \partial^\nu D_F(x-y) = \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} (-(q^0)^2 + \vec{q}^2) \tilde{D}_F(q) = \delta^4(x-y) \tag{4.56}$$

and, since the Fourier transform of the δ function is 1, $-q^2 \tilde{D}_F(q) = 1$. Thus,

$$\tilde{D}_F(q) = \frac{-1}{q^2} \tag{4.57}$$

for $q^2 \neq 0$, and in analogy with the e^- propagator, we avoid the poles at $q^2 = 0$ by writing

$$\tilde{D}_F(q) = \frac{-1}{q^2 + i\epsilon} \tag{4.58}$$

with $\epsilon \to 0^+$. Thus the photon propagator is

$$D_F(x-y) = \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} \frac{-1}{q^2 + i\epsilon}\,, \tag{4.59}$$

and, hence,

$$A^\mu(x) = \int d^4y D_F(x-y) J^\mu(y) \tag{4.60}$$

obeys

$$\partial_\nu \partial^\nu A^\mu(x) = J^\mu(x)\,. \tag{4.61}$$

Now, we can write the S-matrix element

$$S_{fi} = -ie\int d^4x \overline{\psi}_f(x) \not{A}(x)\psi_i(x) = -ie\int d^4x \overline{\psi}_f(x)\gamma_\mu\psi_i(x)A^\mu(x) \tag{4.62}$$

as

$$S_{fi} = -ie\int d^4x \int d^4y \overline{\psi}_f(x)\gamma_\mu\psi_i(x)D_F(x-y)J^\mu(y)\,. \tag{4.63}$$

Next, we need an expression for the proton current $J^\mu(y)$:

Recall that the probability density of a Dirac particle is $\psi^\dagger\psi = \overline{\psi}\gamma^0\psi$ and the probability current is $\psi^\dagger\vec{\alpha}\psi = \overline{\psi}\gamma^0\vec{\alpha}\psi = \overline{\psi}\vec{\gamma}\psi$. For a particle of charge q, this will give the charge density $\rho = q\overline{\psi}\gamma^0\psi$ and the current density vector $\vec{j} = q\overline{\psi}\vec{\gamma}\psi$. Thus the 4-vector current is $J^\mu = q\overline{\psi}\gamma^\mu\psi$. For a proton with initial wavefunction $\psi_i^p(x)$, final wave function $\psi_f^p(x)$ and charge $+e$, the generalization of J^μ is the so called Electromagnetic Transition Current

$$J^\mu = e\overline{\psi}_f^p(x)\gamma^\mu\psi_i^p(x)\,, \tag{4.64}$$

so that

$$S_{fi} = i\int d^4x \int d^4y\big(-e\overline{\psi}_f(x)\gamma_\mu\psi_i(x)\big)D_F(x-y)\big(+e\overline{\psi}_f^p(x)\gamma^\mu\psi_i^p(x)\big)\,. \tag{4.65}$$

The S-matrix element couples the EM transition current for the electron to the EM transition current for the proton through the photon propagator.

Normalising in a box of volume V, as in the last section, the free particle wave functions for the incoming and outgoing electrons and protons are:

$$\begin{aligned}
\psi_i(x) &= \frac{1}{\sqrt{2E_iV}}e^{-ip_i\cdot x}U(p_i,s_i)\\
\psi_f(x) &= \frac{1}{\sqrt{2E_fV}}e^{-ip_f\cdot x}U(p_f,s_f)\\
\psi_i^p(y) &= \frac{1}{\sqrt{2\mathcal{E}_iV}}e^{-iP_i\cdot y}U(P_i,S_i)\\
\psi_f^p(y) &= \frac{1}{\sqrt{2\mathcal{E}_fV}}e^{-iP_f\cdot y}U(P_f,S_f)
\end{aligned} \tag{4.66}$$

where $\mathcal{E}_i$ and $\mathcal{E}_f$ are the proton energies. Therefore,

$$\begin{aligned}
S_{fi} &= -ie^2\frac{1}{\sqrt{2E_iV}}\frac{1}{\sqrt{2E_fV}}\frac{1}{\sqrt{2\mathcal{E}_iV}}\frac{1}{\sqrt{2\mathcal{E}_fV}}\overline{U}(p_f,s_f)\gamma_\mu U(p_i,s_i)\overline{U}(P_f,S_f)\gamma^\mu U(P_i,S_i)\\
&\quad\times\int d^4x\int d^4y\, e^{i(p_f-p_i)\cdot x}e^{i(P_f-P_i)\cdot y}D_F(x-y)\,.
\end{aligned} \tag{4.67}$$

The integral in (4.67) is

$$\begin{aligned} I &= \int d^4x \int d^4y\, e^{i(p_f-p_i)\cdot x} e^{i(P_f-P_i)\cdot y} \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} \frac{-1}{q^2+i\epsilon} \\ &= \int \frac{d^4q}{(2\pi)^4} \frac{-1}{q^2+i\epsilon} (2\pi)^4 \delta^4(q-p_f+p_i) \int d^4y e^{i(q+P_f-P_i)\cdot y} \\ &= \frac{-1}{q^2+i\epsilon} \int d^4y e^{i(p_f-p_i+P_f-P_i)\cdot y} \\ &= \frac{-1}{q^2+i\epsilon} (2\pi)^4 \delta^4(p_f-p_i+P_f-P_i) \end{aligned} \tag{4.68}$$

where q is set to $q = p_f - p_i$ due to the δ function in the second line!

Thus,

$$S_{fi} = \frac{1}{\sqrt{2E_iV}} \frac{1}{\sqrt{2E_fV}} \frac{1}{\sqrt{2\mathcal{E}_iV}} \frac{1}{\sqrt{2\mathcal{E}_fV}} (2\pi)^4 \delta^4(p_f-p_i+P_f-P_i) M_{fi} \tag{4.69}$$

where we have introduced the important quantity M_{fi}, which is also called *Invariant Amplitude* or *Invariant Matrix Element*, and

$$M_{fi} = -\frac{ie^2}{q^2} \overline{U}(p_f, s_f) \gamma_\mu U(p_i, s_i) \overline{U}(P_f, S_f) \gamma^\mu U(P_i, S_i) \tag{4.70}$$

with $q = p_f - p_i$. The Transition Amplitude from state i to state f is given by

$$|S_{fi}|^2 = \frac{|M_{fi}|^2}{16 E_i E_f \mathcal{E}_i \mathcal{E}_f V^4} \left[(2\pi)^4 \delta^4(p_f-p_i+P_f-P_i)\right]^2 . \tag{4.71}$$

Using the interpretation of the square of the δ function as in Section 4.1 we obtain for large times T

$$|S_{fi}|^2 = \frac{|M_{fi}|^2}{16 E_i E_f \mathcal{E}_i \mathcal{E}_f V^4} (2\pi)^4 \delta^4(p_f-p_i+P_f-P_i) VT\,, \tag{4.72}$$

where we got an additional factor of the volume V compared to Section 4.1 because of the square of three δ functions for the three spatial momenta!

4.5 Evaluation of $|M_{fi}|^2$ for Electron-Proton Scattering

As in Section 4.1, we assume that the incident beams of electrons and protons are unpolarised with each spin state equally likely. Hence, we have to average over initial spin states. We also assume that we do not observe the final spin states of the electrons and

protons. Therefore, we also have to sum over the final spin states. Thus, in order to evaluate the cross section we need to evaluate

$$\frac{1}{4}\sum_{s_f,s_i,S_f,S_i}|M_{fi}|^2 \equiv |\overline{M}_{fi}|^2\,, \tag{4.73}$$

where the prefactor 1/4 comes from averaging over the initial electron and proton spin states. In particular we need to calculate

$$\begin{aligned}
&\sum_{s_f,s_i,S_f,S_i}|\overline{U}(p_f,s_f)\gamma_\mu U(p_i,s_i)\overline{U}(P_f,S_f)\gamma^\mu U(P_i,S_i)|^2\\
=&\sum_{s_f,s_i,S_f,S_i}\big[\overline{U}(p_f,s_f)\gamma_\mu U(p_i,s_i)\overline{U}(P_f,S_f)\gamma^\mu U(P_i,S_i)\\
&\quad\times\big(\overline{U}(p_f,s_f)\gamma_\nu U(p_i,s_i)\big)^*\big(\overline{U}(P_f,S_f)\gamma^\nu U(P_i,S_i)\big)^*\big]\\
=&\sum_{s_f,s_i}\overline{U}(p_f,s_f)\gamma_\mu U(p_i,s_i)\overline{U}(p_i,s_i)\underbrace{\overline{\gamma}_\nu}_{=\gamma_\nu}U(p_f,s_f)\\
&\times\sum_{S_f,S_i}\overline{U}(P_f,S_f)\gamma^\mu U(P_i,P_i)\overline{U}(P_i,S_i)\underbrace{\overline{\gamma}^\nu}_{=\gamma^\nu}U(P_f,S_f)\\
=&\;Tr\big((\not{p}_f+m)\gamma_\mu(\not{p}_i+m)\gamma_\nu\big)Tr\big((\not{P}_f+M)\gamma^\mu(\not{P}_i+M)\gamma^\nu\big)\,,
\end{aligned}\tag{4.74}$$

where we have used several identities from Section 4.2.

The trace can be evaluated to obtain

$$Tr\big((\not{p}_f+m)\gamma_\mu(\not{p}_i+m)\gamma_\nu\big)=4\left[p_f^\mu p_i^\nu+p_f^\nu p_i^\mu+g^{\mu\nu}(m^2-p_f\cdot p_i)\right]\tag{4.75}$$

and similarly for the proton trace. Thus,

$$\begin{aligned}
&\frac{1}{16}Tr\big((\not{p}_f+m)\gamma_\mu(\not{p}_i+m)\gamma_\nu\big)Tr\big((\not{P}_f+M)\gamma^\mu(\not{P}_i+M)\gamma^\nu\big)\\
=&\left[p_f^\mu p_i^\nu+p_f^\nu p_i^\mu+g^{\mu\nu}(m^2-p_f\cdot p_i)\right]\left[(P_f)_\mu(P_i)_\nu+(P_f)_\nu(P_i)_\mu+g_{\mu\nu}(M^2-P_f\cdot P_i)\right]\\
=&\;2p_f\cdot P_f p_i\cdot P_i+2p_f\cdot P_i p_i\cdot P_f-2m^2P_f\cdot P_i-2M^2p_f\cdot p_i+4m^2M^2
\end{aligned}\tag{4.76}$$

and finally we obtain

$$\begin{aligned}
|\overline{M}_{fi}|^2 &= \frac{1}{4}\sum_{s_f,s_i,S_f,S_i}|M_{fi}|^2 \qquad (4.77)\\
&= \frac{8e^4}{(q^2)^2}\left[p_f\cdot P_f p_i\cdot P_i+p_f\cdot P_i p_i\cdot P_f-m^2P_f\cdot P_i-M^2p_f\cdot p_i+2m^2M^2\right]\,.
\end{aligned}$$

4.6 $\frac{d\sigma}{d\Omega}$ for Electron-Proton Scattering

The Transistion Amplitude per unit time per unit volume into the range of finite momenta $\vec{p}_f$ to $\vec{p}_f + d\vec{p}_f$ and $\vec{P}_f$ to $\vec{P}_f + d\vec{P}_f$ is

$$\begin{aligned}
&\frac{|S_{fi}|^2}{VT}\frac{V}{(2\pi)^3}d^3p_f\frac{V}{(2\pi)^3}d^3P_f = \frac{|S_{fi}|^2}{T}\frac{V}{(2\pi)^6}d^3p_f d^3P_f \\
&= \frac{|\overline{M}_{fi}|^2}{16E_iE_f\mathcal{E}_i\mathcal{E}_fV^4}(2\pi)^4VT\delta^4(p_f + P_f - p_i - P_i)\frac{V}{T(2\pi)^6}d^3p_f d^3P_f \\
&= \frac{|\overline{M}_{fi}|^2}{16V^2E_iE_f\mathcal{E}_i\mathcal{E}_f(2\pi)^2}\delta^4(p_f + P_f - p_i - P_i)d^3p_f d^3P_f\,. \qquad (4.78)
\end{aligned}$$

To obtain the differential cross section for the scattered electron $\frac{d\sigma_e}{d\Omega_e}$ we need:

$$d\sigma_e = \frac{\text{Probability of } e^- \text{ scattering into solid angle } d\Omega_e \text{ per unit time}}{(\text{Nr. of Target Particles})(\text{Nr. of beam particles incident per unit area per unit time})} \qquad (4.79)$$

In our normalisation

$$(\text{Nr. of Target Particles per unit volume}) = \frac{1}{V}\,, \qquad (4.80)$$

also for collinear beams

$$\begin{aligned}
(\text{Nr. of beam particles incident per unit area per unit time}) &= \frac{1}{V}\left|\vec{v}_i^e - \vec{v}_i^P\right| \\
&= \frac{1}{V}\left|\frac{|\vec{p}_i|}{E_i} + \frac{|\vec{P}_i|}{\mathcal{E}_i}\right| \qquad (4.81)
\end{aligned}$$

To obtain $\frac{d\sigma_e}{d\Omega_e}$ we must integrate d^3P_f and $d|\vec{p}_f|$ but NOT $d\Omega_e$. Now,

$$d^3p_f = |\vec{p}_f|^2 d|\vec{p}_f|d\Omega_e \qquad (4.82)$$

and as in Section 4.1, $|\vec{p}_f|d|\vec{p}_f| = E_f dE_f$. Hence,

$$d^3p_f = |\vec{p}_f|E_f dE_f d\Omega_e\,. \qquad (4.83)$$

The $\int d^3P_f$ integration is trivial because of the δ functions in (4.78), i.e. we simply have to evaluate everything at $\vec{P}_f = \vec{P}_i + \vec{p}_i - \vec{p}_f$. We find

$$d\sigma_e = \int \frac{|\overline{M}_{fi}|^2}{16V^2E_iE_f\mathcal{E}_i\mathcal{E}_f(2\pi)^2}\frac{|\vec{p}_f|E_f dE_f d\Omega_e}{\frac{1}{V^2}\left|\vec{v}_i^e - \vec{v}_i^P\right|}\delta(E_f + \mathcal{E}_f - E_i - \mathcal{E}_i) \qquad (4.84)$$

evaluated at $\vec{P}_f = \vec{P}_i + \vec{p}_i - \vec{p}_f$. Therefore,

$$\frac{d\sigma_e}{d\Omega_e} = \frac{|\overline{M}_{fi}|^2}{16E_i\mathcal{E}_i\mathcal{E}_f(2\pi)^2}\frac{|\vec{p}_f|}{|\vec{v}_i^e - \vec{v}_i^P|} \tag{4.85}$$

evaluated at $P_f = P_i + p_i - p_f$.

For a target proton initially at rest $p_i = (E_i, \vec{p}_i)$, $p_f = (E_f, \vec{p}_f)$, $P_i = (\mathcal{E}_i = M, \vec{P}_i = \vec{0})$. Then,

$$\left|\vec{v}_i^e - \vec{v}_i^P\right| = \frac{\vec{p}_i}{E_i} \tag{4.86}$$

and

$$\frac{d\sigma_e}{d\Omega_e} = \frac{|\overline{M}_{fi}|^2}{16(2\pi)^2\mathcal{E}_i\mathcal{E}_f}\frac{|\vec{p}_f|}{|\vec{p}_i|} \tag{4.87}$$

Slow Electrons: $E_i, E_f << M$

In this case the proton recoil is tiny and $\mathcal{E}_f \sim \mathcal{E}_i = M$, $E_f \sim E_i$, $|\vec{p}_f| \sim |\vec{p}_i|$ and $\vec{P}_f \sim \vec{0}$. This is like scattering off a static nucleus with $Z = 1$ and we should recover the Mott cross section formula.

$$\frac{d\sigma_e}{d\Omega_e} = \frac{|\overline{M}_{fi}|^2}{16(2\pi)^2M^2} \tag{4.88}$$

Also, with $|\overline{M}_{fi}|^2$ from Section 4.6, we obtain

$$|\overline{M}_{fi}|^2 \sim \frac{8e^4}{(q^2)^2}M^2\left(2E_i^2 + m^2 - p_f \cdot p_i\right) \tag{4.89}$$

with, $q = p_f - p_i \sim (0, \vec{p}_f - \vec{p}_i) = (0, \vec{q})$. Using $|\vec{q}|^4 = 16|\vec{p}_i|^4\sin^4(\theta/2)$ from Section 4.3 and after further simplifications we recover the Mott cross section (with $Z = 1$)

$$\frac{d\sigma_e}{d\Omega_e} = \frac{\alpha^2}{4|\vec{p}_i|^2\beta^2\sin^4(\theta/2)}\left(1 - \beta^2\sin^2(\theta/2)\right) \tag{4.90}$$

where $\beta \equiv |\vec{p}_i|/E_i$.

Ultrarelativistic Electrons: $E_i, E_f >> m$ where m is the electron mass. In this case simplification of $|\overline{M}_{fi}|^2$ leads to

$$\frac{d\sigma_e}{d\Omega_e} = \frac{\alpha^2}{4E_i^2}\frac{\left[\cos^2\frac{\theta}{2} - \frac{q^2}{2M^2}\sin^2\frac{\theta}{2}\right]}{\sin^4\frac{\theta}{2}\left[1 + \frac{2E_i}{M}\sin^2\frac{\theta}{2}\right]} \tag{4.91}$$

which agrees with experiment.

4.7 Feynman Rules

More generally, the Invarian Amplitude M_{fi} derived from a Feynman diagram is given by the following *Rules*:

(a) Associate with each incoming/outgoing external electron (or spin-$\frac{1}{2}$ particle) line of 4-momentum p_i/p_f a factor $U(p_i, s_i)/\overline{U}(p_f, s_f)$.

(b) Associate with each incoming/outgoing negative energy electron line a factor $V(p_i, s_i)/\overline{V}(p_f, s_f)$ where $-p_i$ / $-p_f$ are the 4-momenta of the negative energy electrons, i.e. p_i/p_f are the 4-momenta of the positive energy positrons.

(c) Associate with each internal photon line of 4-momentum q a factor $-i\frac{g_{\mu\nu}}{q^2}$.

(d) Associate with each interaction vertex a factor $-ie\gamma^\mu$ for a particle (or negative energy particle) of charge e.

The Scattering Amplitude S_{fi} is related to the Invariant Amplitude M_{fi} as in Section 4.5. The differential cross section can then be calculated from $|S_{fi}|^2$ using the methods developed earlier.

Example 1: For electron-positron scattering there are two Feynman diagrams contributing to M_{fi}

(picture goes here)

Then,

$$\begin{aligned} S_{fi} &= \frac{(2\pi)^4\delta^4(p_3+p_4-p_1-p_2)}{\sqrt{2E_1V}\sqrt{2E_2V}\sqrt{2E_3V}\sqrt{2E_4V}}\left[\frac{-i}{(p_3-p_1)^2}\overline{U}(p_3)(-ie\gamma_\mu)U(p_1)\overline{V}(p_2)(-ie\gamma^\mu)V(p_4)\right. \\ &- \left.\frac{-i}{(p_1+p_2)^2}\overline{V}(p_2)(-ie\gamma_\mu)U(p_1)\overline{U}(p_3)(-ie\gamma^\mu)V(p_4)\right] \end{aligned} \tag{4.92}$$

Note that the relative minus sign between the first and second term has been introduced to ensure ANTI-SYMMETRY under the interchange of the two incoming electrons (one of which has negative energy) and under the interchange of the two outgoing electrons. Note that this anti-symmetrisation is only necessary if identical particles, in this case electrons, are considered:

$U(p_1) \leftrightarrow V(p_4)$ OR $\overline{U}(p_3) \leftrightarrow \overline{V}(p_2)$

Example 2: Electron scattering off a Dirac proton. There is only one diagram to consider (photon exchange between the two particles) because the two particles are not identical. It

is left as an exercise to check that the above stated Feynman reproduce the same invariant matrix element M_{fi} and, hence, the same scattering amplitude as derived in Section 4.4.

Example 3: Electron-electron scattering. As in Example 1 there are two Feynman diagrams to consider; again with a relative minus sign between the two terms because of the anti-symmetrisation under the exchange of the two incoming (and outgoing) electrons. Denote the incoming electrons by 1 and 2, and the outgoing ones by 3 and 4, hence the amplitude must be anti-symmetric under the exchange $1 \leftrightarrow 2$ and the exchange $3 \leftrightarrow 4$. Application of the Feynman rules yields:

$$\begin{aligned} S_{fi} &= \frac{(2\pi)^4\delta^4(p_3+p_4-p_1-p_2)}{\sqrt{2E_1V}\sqrt{2E_2V}\sqrt{2E_3V}\sqrt{2E_4V}}\left[\frac{-i}{(p_3-p_1)^2}\overline{U}(p_3)(-ie\gamma_\mu)U(p_1)\overline{U}(p_4)(-ie\gamma^\mu)U(p_2)\right. \\ &- \left.\frac{-i}{(p_1-p_4)^2}\overline{U}(p_4)(-ie\gamma_\mu)U(p_1)\overline{U}(p_3)(-ie\gamma^\mu)U(p_2)\right] \end{aligned} \tag{4.93}$$

(picture goes here)

5 Relativistic Quantum Field Theory

5.1 Background

The first example of a classical field theory was electromagnetism as described by Maxwell

$$\partial^\mu F_{\mu\nu} = j_\nu \ , \ \ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \, . \tag{5.1}$$

The first Quantum Field Theory (QFT) was the theory of the electromagnetic field $A_\mu(x)$. Quantisation of the electromagnetic field by imposing commutation relations led to a particle interpretation in terms of photons. The quantised electromagnetic field A_μ becomes an operator that can create and annihilate photons. In simple terms, the quantised field can be viewed as an infinite collection of harmonic oscillators.

To begin with let us consider the simpler case of the Klein-Gordon field $\phi(x)$. The main novelty here is that ϕ is treated as an operator rather than a wave function as we did in RQM.

5.2 Field Equations

To a field $\phi(x)$ we assign a Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$, and the corresponding Lagrangian is

$$L = \int d^3x \mathcal{L}(\phi, \partial_\mu \phi) \,. \tag{5.2}$$

The action is then defined as

$$I = \int_{t_1}^{t_2} dt L = \int_{t_1}^{t_2} dt \int d^3x \mathcal{L}(\phi, \partial_\mu \phi) \,. \tag{5.3}$$

The field equations or Euler-Lagrange equations are given by Hamilton's principle or variational principle $\delta I = 0$ for arbitrary variations $\delta\phi$ of the field, with $\delta\phi = 0$ at the boundaries $t = t_1$ and $t = t_2$. For arbitrary variations of ϕ

$$\begin{aligned}
\delta I &= \delta(\int_{t_1}^{t_2} dt L) = \delta(\int_{t_1}^{t_2} dt \int d^3x \mathcal{L}(\phi, \partial_\mu \phi)) \\
&= \int_{t_1}^{t_2} dt \int d^3x \left[\mathcal{L}(\phi + \delta\phi, \partial_\mu \phi + \partial_\mu(\delta\phi)) - \mathcal{L}(\phi, \partial_\mu \phi)\right] \\
&= \int_{t_1}^{t_2} dt \int d^3x \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\mu(\delta\phi)\right] \\
&= \int_{t_1}^{t_2} dt \int d^3x \delta\phi \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right] = 0 \,,
\end{aligned} \tag{5.4}$$

where we dropped a boundary term coming from partial integration due to the boundary condition on $\delta\phi$. Since this must be true for arbitrary variation $\delta\phi$ we arrive at the field equations

$$\frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \,. \tag{5.5}$$

5.3 The Neutral Klein-Gordon Field

Consider a Hermitian field operator $\phi(x) = \phi^\dagger(x)$ which satisfies the Klein-Gordon equation

$$(\partial_\mu \partial^\mu + m^2)\phi = 0 \,. \tag{5.6}$$

Note that ϕ is an operator and NOT a wave function (which would not make much sense since ϕ is Hermitian i.e. real now). The Klein-Gordon equation is the field equation derived from the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2\right) \,. \tag{5.7}$$

Any solution of the KG equation can be decomposed in terms of positive and negative energy plane wave solutions (complete set of eigenfunctions)

$$\phi(x) = \int d^3k \left(a_+(k) f_k^{(+)}(x) + a_-(k) f_k^{(-)}(x) \right) \tag{5.8}$$

where $f_k^{\pm}(x) = \frac{1}{\sqrt{(2\pi)^3 2E_k}} e^{\mp ik\cdot x}$ with $E_k = +\sqrt{\vec{k}^2 + m^2}$. If ϕ is Hermitian, $\phi^\dagger = \phi$, then $a_-(k) = a_+^\dagger(k)$. Write $a_+(k) \equiv a(k)$, $a_-(k) \equiv a^\dagger(k)$. Hence,

$$\phi(x) = \int d^3k \left(a(\vec{k}) f_k^{(+)}(x) + a^\dagger(\vec{k}) f_k^{(-)}(x) \right) . \tag{5.9}$$

It can be checked that $f_k^{\pm}(x)$ obey the orthonormality conditions

$$\begin{aligned} i \int d^3x f_k^{(\pm)}(x) \overleftrightarrow{\partial_0} f_k^{(\pm)}(x) &= 0 \\ i \int d^3x f_k^{(\mp)}(x) \overleftrightarrow{\partial_0} f_k^{(\pm)}(x) &= \pm \delta^3(\vec{k} - \vec{k}') \end{aligned} \tag{5.10}$$

where $A \overleftrightarrow{\partial_0} B = A\partial_0 B - (\partial_0 A)B$. With the help of these orthonormality conditions we can calculate a and $a^\dagger$ in terms of ϕ.

$$\begin{aligned} & \int d^3x f_{k'}^{(+)*}(x) \overleftrightarrow{\partial_0} \phi(x) \\ = & \int d^3k \left(a(k) \int d^3x f_{k'}^{(+)*}(x) \overleftrightarrow{\partial_0} f_k^{(+)}(x) \right. \\ & \left. + a^\dagger(k) \int d^3x f_{k'}^{(+)*}(x) \overleftrightarrow{\partial_0} f_k^{(-)}(x) \right) \\ = & \int d^3k\, a(k)(-i\delta^3(\vec{k'} - \vec{k})) = -ia(k') . \end{aligned} \tag{5.11}$$

Hence,

$$a(k) = i \int d^3x f_k^{(+)*}(x) \overleftrightarrow{\partial_0} \phi(x) \tag{5.12}$$

and similary by taking the Hermitian conjugate we find

$$a^\dagger(k) = -i \int d^3x f_k^{(-)*}(x) \overleftrightarrow{\partial_0} \phi(x) . \tag{5.13}$$

The theory is quantised by first introducing a generalised momentum (density) Π conjugate to ϕ by

$$\Pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \partial_0 \phi = \dot{\phi} . \tag{5.14}$$

In analogy to the QM commutation relation $[p, x] = -i\hbar$ we impose the canonical commutation relations

$$\begin{aligned}
&[\Pi(t,\vec{x}), \phi(t,\vec{x}')] = -i\delta^3(\vec{x}-\vec{x}') = \left[\dot{\phi}(t,\vec{x}), \phi(t,\vec{x}')\right] \\
&[\phi(t,\vec{x}), \phi(t,\vec{x}')] = [\Pi(t,\vec{x}), \Pi(t,\vec{x}')] = 0\,.
\end{aligned} \tag{5.15}$$

The corresponding commutation relations for $a(k)$ and $a^\dagger(k)$ can now be derived

$$\begin{aligned}
&\left[a(k), a^\dagger(k')\right] = i(-i)\int d^3x \int d^3y \left[f_k^{(+)*}(x) \overleftrightarrow{\partial_0} \phi(x), f_{k'}^{(-)*}(y) \overleftrightarrow{\partial_0} \phi(y)\right] \\
&= \int d^3x \int d^3y \left[f_k^{(+)*}(x)\frac{\partial}{\partial x_0}\phi(x) - \frac{\partial}{\partial x_0} f_k^{(+)*}(x)\phi(x), f_{k'}^{(-)*}(y)\frac{\partial}{\partial y_0}\phi(y) - \frac{\partial}{\partial y_0} f_{k'}^{(-)*}(y)\phi(y)\right] \\
&= \int d^3x \int d^3y \left[-f_k^{(+)*}(x) \underbrace{[\dot{\phi}(x), \phi(y)]}_{=-i\delta^3(\vec{x}-\vec{y})} \frac{\partial}{\partial y_0} f_{k'}^{(-)*}(y) - \frac{\partial}{\partial x_0} f_k^{(+)*}(x) \underbrace{[\phi(x), \dot{\phi}(y)]}_{=+i\delta^3(\vec{y}-\vec{x})} f_{k'}^{(-)*}(y)\right] \\
&= i\left[\int d^3x f_k^{(+)*}(x)\frac{\partial}{\partial x_0} f_{k'}^{(-)*}(x) - \int d^3x \frac{\partial}{\partial x_0} f_k^{(+)*}(x) f_{k'}^{(-)*}(x)\right] \\
&= i\int d^3x f_k^{(+)*}(x) \overleftrightarrow{\partial_0} f_{k'}^{(-)*}(x) \\
&= \delta^3(\vec{k}-\vec{k}')\,,
\end{aligned} \tag{5.16}$$

hence,

$$\left[a(k), a^\dagger(k')\right] = \delta^3(\vec{k}-\vec{k}')\,. \tag{5.17}$$

Similarly,

$$[a(k), a(k')] = \left[a^\dagger(k), a^\dagger(k')\right] = 0\,. \tag{5.18}$$

Now write the Hamiltonian in terms of a and $a^\dagger$. Begin with the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left((\dot{\phi})^2 - (\vec{\nabla}\phi)^2 - m^2\phi^2\right)\,. \tag{5.19}$$

Compare this with the Lagrangian $L = T - V$ and Hamiltonian $H = T + V$ in classical mechanics. Therefore the Hamiltonian density is

$$\mathcal{H} = \frac{1}{2}\left((\dot{\phi})^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2\right)\,. \tag{5.20}$$

and the Hamiltonian is $H = \int d^3x \mathcal{H}$. Now,

$$\begin{aligned}
\dot{\phi} &= -i \int d^3k k_0 \left(f_k^{(+)}(x) a(k) - f_k^{(-)}(x) a^\dagger(k) \right) \\
\vec{\nabla}\phi &= i \int d^3k \vec{k} \left(f_k^{(+)}(x) a(k) - f_k^{(-)}(x) a^\dagger(k) \right) \\
H &= \frac{1}{2} \int d^3x \frac{1}{2} \left((\dot{\phi})^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2 \right) \\
&= \int d^3k \int d^3k' \int d^3x (-k_0 k_0' - \vec{k}\cdot\vec{k}') \times \left(f_k^{(+)} f_{k'}^{(+)} a(k) a(k') \right. \\
&\quad + f_k^{(-)} f_{k'}^{(-)} a^\dagger(k) a^\dagger(k') - f_k^{(+)} f_{k'}^{(-)} a(k) a^\dagger(k') - f_k^{(-)} f_{k'}^{(+)} a^\dagger(k) a(k') \Big) \\
&\quad + \int d^3k \int d^3k' \int d^3x m^2 \times \left(f_k^{(+)} f_{k'}^{(+)} a(k) a(k') \right. \\
&\quad \left. + f_k^{(-)} f_{k'}^{(-)} a^\dagger(k) a^\dagger(k') + f_k^{(+)} f_{k'}^{(-)} a(k) a^\dagger(k') + f_k^{(-)} f_{k'}^{(+)} a^\dagger(k) a(k') \right) \qquad (5.21)
\end{aligned}$$

Using the identities

$$\begin{aligned}
\int d^3x f_k^{(\pm)}(x) f_{k'}^{(\pm)}(x) &= \frac{1}{2E_k} e^{\mp 2iE_k x^0} \delta^3(\vec{k} + \vec{k}') \\
\int d^3x f_k^{(\pm)}(x) f_{k'}^{(\mp)}(x) &= \frac{1}{2E_k} \delta^3(\vec{k} - \vec{k}') \qquad (5.22)
\end{aligned}$$

where $E_k = +\sqrt{\vec{k}^2 + m^2}$, we obtain

$$\begin{aligned}
H &= \int d^3k \frac{1}{4E_k} \left(a(k) a^\dagger(k) + a^\dagger(k) a(k) \right) \underbrace{(k_0^2 + \vec{k}^2 + m^2)}_{2E_k^2} \\
&= \int d^3k \frac{E_k}{2} \left(a(k) a^\dagger(k) + a^\dagger(k) a(k) \right) . \qquad (5.23)
\end{aligned}$$

It is straightforward to check that

$$\begin{aligned}
[H, a(k)] &= -E_k a(k) \\
\left[H, a^\dagger(k)\right] &= -E_k a^\dagger(k) . \qquad (5.24)
\end{aligned}$$

Consequently, a and $a^\dagger$ behave as the lowering and raising operators for the harmonic oscillator! The vacuum state $|0\rangle$ is defined as the state for which $a(k)|0\rangle = 0$ for all k.

Energy of the Vacuum

$$\begin{aligned}
H &= \int d^3k \frac{E_k}{2} (a(k) a^\dagger(k) + a^\dagger(k) a(k)) \\
&= \int d^3k \frac{E_k}{2} (2a^{\dagger k} a(k) + \delta^3(\vec{0})) \\
&= \int d^3k E_k a^{\dagger k} a(k) + \frac{1}{2} \int d^3k E_k \delta^3(\vec{0}) \qquad (5.25)
\end{aligned}$$

The first term acting on the vacuum gives zero, but the second term yields an infinite vacuum energy. Since we are only interested in differences between energies, we subtract off the infinite vacuum energy.

Thus, we replace the Hamiltonian H by its normal ordered counterpart $\tilde{H}$, which is defined as

$$\tilde{H} = \int d^3k E_k a^{\dagger k} a(k)\,, \tag{5.26}$$

and leads to a vanishing vacuum energy

$$\langle 0|\tilde{H}|0\rangle = 0\,, \tag{5.27}$$

because of the definition of the vacuum state!

The state obtained by m applications of $a^\dagger(k)$ on the vacuum state has energy

$$E_m = mE_k\,. \tag{5.28}$$

Therefore, we may interpret $a^\dagger$ ($a(k)$) as operators creating (annihilating) quanta/particles with four-momentum k.

5.4 Interactions

Also the Dirac field ψ can be quantised, now using anti-commutation relations instead of canonical commutation relations. The free Lagrangian density for Dirac particles and photons turns out to be

$$\mathcal{L}_{free} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{5.29}$$

Introducing the electromagnetic interaction by the usual "minimal substitution"

$$\partial_\mu \to \partial_\mu + ieA_\mu \tag{5.30}$$

we get

$$\mathcal{L} = \mathcal{L}_{free} + \mathcal{L}_I = \overline{\psi}(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, \tag{5.31}$$

where the interaction Lagrangian is given by

$$\mathcal{L}_I = e\overline{\psi}\gamma^\mu A_\mu\psi\,. \tag{5.32}$$

The detailed development of QFT leads in the end to the same Feynman rules as for RQM. Notice the resemblance to $S_{fi} = -ie\int d^4x\overline{\psi}_f(x)\gamma^\mu A_\mu(x)\psi_i(x)$.

References

[1] J. Bjorken and S. Drell, "Relativistic Quantum Mechanics" and "Relativistic Quantum Fields", McGraw-Hill.

[2] M.E. Peskin and D.V. Schroeder, "An Introduction to Quantum Field Theory", Addison-Wesley.

[3] S. Weinberg, "The Quantum Theory of Fields", Cambridge University Press.

[4] C. Itzykson and J.-B. Zuber, "Quantum Field Theory", McGraw-Hill.

[5] I.J. Aitchison and A.J. Hey, "Gauge Theories in Particle Physics", Adam Hilger.

[6] F. Mandl, "Quantum Mechanics", J Wiley.

[7] W. Greiner, "Relativistic Quantum Mechanics: Wave Equations", Springer-Verlag.