Farrukh Azfar

Publishers' page

Publishers' page

Publishers' page

Publishers’ page

Contents

Farrukh Azfar

Chapter 1

A review of selected topics in special relativity and electromagnetism

TEST In this first chapter we'll quickly review some selected topics in special relativity and electromagnetism is presented. The goal is to familiarize the student with the 4-vector notation of special relativity and identify 4-vectors and Lorentz invariants. The idea is to remind the student how several observables change from one frame to another and how their transformation properties can be generalized. First introduced are contravariant and covariant vectors and their transformation properties. Invariant quantities constructed from these are also defined and the notation used in further chapters is introduced. The 4-vectors include the 4-velocity and 4-momentum of a particle, the charge and current densities, and the vector and scalar potentials. Next the modification of the 4-momentum of a charged particle in the presence of an electro-magnetic field is justified, this will be used later to introduce the electromagnetic field as a perturbation to the Dirac equation as a way of treating scattering problems. In this chapter Gaussian-CGS [1] units have been used throughout and at the end of this chapter natural units are introduced. This chapter contains a lot of material originating from answers given to students of varying backgrounds who had either not covered some material or required a refresher-this leads to a somewhat ad-hoc format. The derivations presented are almost exactly those presented in response to a student's questions and hence in places can be quite detailed or depending on the particular students background. Particular attention is paid to developing the notation that will be used in later chapters.

1.1 A quick review of Lorentz transformations

We begin by writing down a simple Lorentz transformation, the primed (x', y', z', t') and unprimed (x, y, z, t) are the co-ordinates of an event seen in two frames, with the primed frame taken to move to the right along the $+x$ direction relative to the unprimed frame. The Lorentz transformation equations relating the co-ordinates of an event viewed by both observers is then. Defining $\beta = \frac{v}{c}$, $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ we write down the Lorentz transformation relating the space and time co-ordinates of the two frames

$$\begin{aligned} ct' &= \frac{ct - \beta x}{\sqrt{1-\beta^2}} = \gamma(ct - \beta x) \\ x' &= \frac{x - \beta(ct)}{\sqrt{1-\beta^2}} = \gamma(x - \beta(ct)) \\ y' &= y \\ z' &= z \end{aligned} \tag{1.1}$$

We note that $\beta = \vec{\beta}$ and $\gamma = \frac{1}{\sqrt{1-\frac{\vec{v}\cdot\vec{v}}{c^2}}}$ but for simplicities sake we only use one component of the velocity. We can represent these transformations in matrix form with the column vector representing the space and time co-ordinates of the event in the relevant frame.

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.2}$$

We now define s^2, where the space and time co-ordinates are intended to represent the difference between those of a single event, so s^2 is the an interval

$$s^2 = c^2 t'^2 - x'^2 - y'^2 - z'^2$$

using Eqn. 1.1 and $\gamma^2 = \frac{1}{1-\beta^2}$ we have $\beta^2(x^2) - x^2 = \frac{-x^2}{\gamma^2}$ we can easily show that

$$s^2 = c^2 t'^2 - x'^2 - y'^2 - z'^2 = c^2 t^2 - x^2 - y^2 - z^2$$

. Thus s^2 is the same in *all* frames of reference. It is our first, very simple, example of a Lorentz-invariant, or a quantity that remains the same from one Lorentz transformation to another. The Lorenz transformation is also referred a to as a Lorentz boost in the $+x$ direction, we will use this term interchangeably in this book. A choice of boost in the $+x$ direction has been made for simplicity's sake. This is an arbitrary choice of axis, and it is easy to see that the invariant remains so even if a velocity in an arbitrary direction is taken for the boost.

The simple boost along $+x$ just described is not the only one that leaves s^2 unchanged, it should be clear that the following two transformations corresponding to inversion of all the space co-ordinates and then an inversion of time-both represented in matrix format-also leave s^2 unchanged.

$$\begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.3}$$

$$\begin{pmatrix} -ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.4}$$

Note that the determinant of the transformation matrices in Eqns. 1.3-1.4 is $= -1$ and the determinant of the Lorentz transformation in the boost along x in Eqn. 1.2 is easily calculated to be $\gamma^2 - \gamma^2\beta^2 = \gamma^2(1-\beta^2) = \frac{1-\beta^2}{1-\beta^2} = 1$. We refer to Lorentz transformations whose determinants are $= -1$ as improper transformations and those with determinant 1 as proper transformations and note that there is no way of going from one to the other via a continuous series of successive boosts, this is our first example of a discreet transformation.

1.2 A different parameterization of the Lorentz Transformation, comparision with rotations

The Lorentz transformations can also be parameterized in terms of hyperbolic functions . We can easily rewrite the relation $\gamma^2 = \frac{1}{1-\beta^2}$ as $\gamma^2 - \gamma^2\beta^2 = 1$ and then note that the two relations

$$\gamma^2 - \gamma^2\beta^2 = 1$$

and

$$\frac{\gamma\beta}{\gamma} = \beta$$

are reminiscent of the following well known relations between the hyperbolic functions :

$$\cosh^2 \chi - \sinh^2 \chi = 1$$

and

$$\frac{\sinh \chi}{\cosh \chi} = \tanh \chi$$

. Using these relations the Lorentz transformations can then be easily reparameterized with the following definitions of γ and β :

$$\begin{aligned} \cosh \chi &= \gamma \\ \sinh \chi &= \gamma\beta \\ \tanh \chi &= \beta \end{aligned} \tag{1.5}$$

The Lorentz transformation in Eqn. 1.2 in matrix form can then be rewritten as:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh \chi & -\sinh \chi & 0 & 0 \\ -\sinh \chi & \cosh \chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.6}$$

It is illustrative to compare this with the matrix equation for the rotation of the x-y plane about the z axis by an angle θ, the rotated (x', y', z') are related to the unrotated (x, y, z) co-ordinates of a point by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1.7}$$

It is trivial to show norm of the vector from the origin to this arbitrary point in both frames, is invariant that is $r^2 = x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2$. Lorentz transformations can therefore be thought of as "rotations" in 4-dimensional space (ct, x, y, z) that leave the "norm" $ct^2 - x^2 - y^2 - z^2$ invariant just as rotations leave $x^2 + y^2 + z^2$ unchanged. The unchanged norm is simply the scalar product $\vec{x} \cdot \vec{x}$, with $\vec{x} = x\hat{x} + y\hat{y} + z\hat{z}$ we can also write this scalar product in the following way:

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1.8}$$

Reminding ourselves of the quantity $c^2t^2 - x^2 - y^2 - z^2$ doesn't change under Lorentz transformations we can also define a scalar product using (ct, x, y, z) grouped together as one vector. This is our first example of a Lorentz vector or a 4-vector) and so :

$$c^2t^2 - x^2 - y^2 - z^2 = \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.9}$$

We have put in the matrix between the two column vectors to introduce the concept of a metric which allows us to calculate an invariant measure(s^2) which is a property of the transformation. The metric used to form the Lorentz invariant scalar product is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and is simply the indentity matrix in equation 1.8. The metric matrix above, used for forming the invariant Lorentz scalar product will be referred to as the metric tensor $g^{\mu\nu}$ or $g_{\mu\nu}$ depending on how it is used, this will become clearer in the next section.

1.3 Covariant and contravariant 4-vectors

There is another way of viewing the process of constructing Lorentz invariants , we define two types of Lorentz 4-vectors :

Let x^μ denote (x^0, x^1, x^2, x^3) where $x^0 = ct$, $x^1 = x$, $x^2 = y$ and $x^3 = z$ (the raised index is *not* a power!) and let x_μ denote (x_0, x_1, x_2, x_3) where $x_0 = ct$ $(= x^0)$, $x_1 = -x$, $x_2 = -y$ and $x_3 = -z$ then the scalar product that generates the Lorentz invariant is simply the product of the components in identical slots of these two different types of vectors.

Thus for the scalar product formed using our spatial vector we have:

$$\sum_{i=1}^{3} x^i x^i = x^2 + y^2 + z^2 \tag{1.10}$$

and for the scalar product formed using our two 4-vectors we have:

$$\sum_{\mu=0}^{3} x_\mu x^\mu = c^2t^2 - x^2 - y^2 - z^2 \tag{1.11}$$

with the following definitions for the two types of 4-vectors:

$$\begin{aligned} x^\mu &= (x^0, x^1, x^2, x^3) = (ct, x, y, z) \\ x_\mu &= (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) \end{aligned} \tag{1.12}$$

The 4-vector represented by the raised index is called a contravariant vector and the one with the lowered index is called a covariant vector . The quantity

$$\sum_{\mu=0}^{3} x_\mu x^\mu = c^2t^2 - x^2 - y^2 - z^2$$

is a Lorentz invariant or a Lorentz scalar ; often just the term scalar or invariant may be used. Greek indices are customarily assumed to run over the 4 values 0, 1, 2, 3 and Roman indices over the spatial 1, 2, 3. Using the metric $g^{\mu\nu}$ we can write:

$$c^2t^2 - x^2 - y^2 - z^2 = \sum_{\nu=0}^{3}\sum_{\mu=0}^{3} g^{\mu\nu} x_\nu x_\mu \tag{1.13}$$

since :

$$\sum_{\nu=0}^{3} g^{\mu\nu} x_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = x^\mu$$

Since we have seen that the operation of $g^{\mu\nu}$ on x_ν changes it to x^μ we can define

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

and the following equations using $g^{\mu\nu}$ and $g_{\mu\nu}$ to change one type of vector to the other are trivial to verify:

$$
\begin{aligned}
x^\mu &= \sum_{\nu=0}^{3} g^{\mu\nu} x_\nu \\
x_\mu &= \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu
\end{aligned}
\tag{1.14}
$$

It should be clear in equation 1.14 that we are multiplying each νth element of the column vector x with each νth element of the μth row of the metric, which in turn is the μth element of the resulting vector.

The choice of naming the indices which represent the sum over a set of elements is arbitrary. Note that we could just as well have written :

$$\sum_{\alpha=0}^{3} g^{\mu\alpha} x_\alpha = x^\mu$$

and

$$\sum_{\beta=0}^{3} g_{\alpha\beta} x^\beta = x_\alpha$$

and the meaning and result are identical. Note the convention of raised and lowered indices in going from one type of vector to the other and also note that the repeated index is summed over–this makes the use of the summation sign redundant and we will do away with it from now on. This means that repeated indices are summed over (Einstein summation convention).

We briefly summarize some things covered so far. First the two types of 4-vectors and how to transform from one to the other, and their definitions

$$
\begin{aligned}
x^\mu = g^{\mu\nu} x_\nu \quad &= (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \vec{x}) \\
x_\alpha = g_{\alpha\beta} x^\beta \quad &= (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) = (ct, -\vec{x})
\end{aligned}
\tag{1.15}
$$

Since one index is repeated this corresponds to a sum over the elements of the metric tensor and the co-variant or contravariant vector, the index that is not repeated can take on all values from 0-3 and this results in *all* components of the resulting contravariant or co-variant vector. The reader is reminded that the invariant formed can be written as:

$$
\begin{aligned}
&g_{\mu\nu} x^\nu x^\mu = x_\mu x^\mu = x_0 x^0 + (-x)x + (-y)y + (-z)(z) \\
&= c^2 t^2 - x^2 - y^2 - z^2
\end{aligned}
\tag{1.16}
$$

Once again we can see that in Eqn. 1.16 there are no free indices (each index is repeated) and the result is a single number-a Lorentz scalar. As an example $g_{\mu\nu}g^{\mu\nu}$ has no free indices and does not represent the multiplication of two matrices but the sum of the products of all entries with themselves and is equal to 4. On the other hand $g_{\mu\alpha}g^{\alpha\nu}$ has two free indices, the $\mu\nu^{th}$ element of which is a sum over all possible alphas. This can most simply be represented as matrix multiplication i.e. one writes the metric tensor as a 4×4 matrix and multiplies it with itself :

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

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = g^{\nu}{}_{\mu} \text{ or } \delta^{\nu}{}_{\mu}$$

where

$$g^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq n \end{cases}$$

Note that g^{ν}_{μ} is *not* $g^{\mu\nu}$ or $g_{\mu\nu}$ it is in fact the identity matrix another expression for which is $\mathbb{I}$.

1.4 Covariant and Contra-variant 4-vectors and their transformation properties

As a reminder Lorentz transformations can be trivially rewritten in terms of the contravariant and covariant vectors as:

$$\begin{aligned} x^{0\prime} &= \gamma(x^0 - \beta x^1) \\ x^{1\prime} &= \gamma(x^1 - \beta x^0) \\ x^{2\prime} &= x^2 \\ x^{3\prime} &= x^3 \end{aligned} \tag{1.17}$$

and

$$\begin{aligned} x_0' &= \gamma(x^0 - \beta x^1) = \gamma x_0 + \beta\gamma x_1 \\ x_1' &= \gamma(-x^1 + \beta\gamma x^0) = \gamma x_1 + \beta\gamma x_0 \\ x_2' &= -x^2 = x_2 \\ x_3' &= -x^3 = x_3 \end{aligned} \tag{1.18}$$

where $x_0 = x^0$, $x_1 = -x^1$, $x_2 = -x^2$ and $x_3 = -x^3$ have been used. The Lorentz transformation matrices for the transformations for the x^μ and x_μ can be easily read off from Eqns. 1.18 and 1.17

$$\begin{pmatrix} x^{0\prime} \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{1.19}$$

and

$$\begin{pmatrix} x_0' \\ x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} \gamma & +\gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{1.20}$$

It is easy to verify that:

$$\begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & +\gamma\beta & 0 & 0 \\ +\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.21}$$

which means that the transformation matrices for x^μ and x_μ (Λ^α_μ and $(\Lambda^{-1})^\alpha_\mu$ are inverses of each other. Thus x^μ transforms according to the Lorentz transformation and x_μ according to the corresponding *inverse* of the same Lorentz transformation. Now any linear transformation such as the Lorentz transformation can be expressed in terms of partial derivatives of the components of one co–ordinate system with respect to the other, so

$$x^{\alpha\prime} = \frac{\partial x^{\alpha\prime}}{\partial x^\beta} x^\beta \tag{1.22}$$

We now define the matrix of Lorentz transformations relating one frame to another is simply the array of partial derivatives $\frac{\partial x^{\alpha\prime}}{\partial x^\beta}$ is defined as:

$$\Lambda^{\alpha}{}_{\beta} = \frac{\partial x^{\alpha\prime}}{\partial x^{\beta}} \tag{1.23}$$

The reason we have used a lower index for the second index of Λ will become apparent when we derive the transformation property of the gradient operator $\frac{\partial}{\partial x^{\alpha}}$ it will turn out that it's transformation property is identical to that of the co-variant vector like x_{α} (and not x^{α}!), this is done in section 1.5. The transformation in equation 1.22 can be expressed as a matrix equation which is written below explicitly:

$$\begin{pmatrix} x^{0\prime} \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{pmatrix} = \begin{pmatrix} \frac{\partial x^{0\prime}}{\partial x^{0}} & \frac{\partial x^{0\prime}}{\partial x^{1}} & \frac{\partial x^{0\prime}}{\partial x^{2}} & \frac{\partial x^{0\prime}}{\partial x^{3}} \\ \frac{\partial x^{1\prime}}{\partial x^{0}} & \frac{\partial x^{1\prime}}{\partial x^{1}} & \frac{\partial x^{1\prime}}{\partial x^{2}} & \frac{\partial x^{1\prime}}{\partial x^{3}} \\ \frac{\partial x^{2\prime}}{\partial x^{0}} & \frac{\partial x^{2\prime}}{\partial x^{1}} & \frac{\partial x^{2\prime}}{\partial x^{2}} & \frac{\partial x^{2\prime}}{\partial x^{3}} \\ \frac{\partial x^{3\prime}}{\partial x^{0}} & \frac{\partial x^{3\prime}}{\partial x^{1}} & \frac{\partial x^{3\prime}}{\partial x^{2}} & \frac{\partial x^{3\prime}}{\partial x^{3}} \end{pmatrix} \begin{pmatrix} x^{0} \\ x^{1} \\ x^{2} \\ x^{3} \end{pmatrix}$$

For the simple case of a boost along x, the horrible matrix above has only the following non-zero entries: $\frac{\partial x^{0\prime}}{\partial x^{0}} = \frac{\partial x^{1\prime}}{\partial x^{1}} = \gamma$, $\frac{\partial x^{0\prime}}{\partial x^{1}} = \frac{\partial x^{1\prime}}{\partial x^{0}} = -\gamma\beta$ and $\frac{\partial x^{2\prime}}{\partial x^{2}} = \frac{\partial x^{3\prime}}{\partial x^{3}} = 1$. Thus our Lorentz transformation writing in only the non-zero terms is:

$$\begin{aligned} x^{0\prime} &= \frac{\partial x^{0\prime}}{\partial x^{0}_{0}} x^{0} + \frac{\partial x^{0\prime}}{\partial x^{1}} x^{1} = \gamma x^{0} - \gamma\beta x^{1} \\ x^{1\prime} &= \frac{\partial x^{1\prime}}{\partial x^{1}} x^{1} + \frac{\partial x^{1\prime}}{\partial x^{0}} x^{0} = \gamma x^{1} - \gamma\beta x^{0} \\ x^{2\prime} &= \frac{\partial x^{2\prime}}{\partial x^{2}} x^{2} = x^{2} \end{aligned}$$

which is of course more compactly expressed as

$$x^{\mu\prime} = \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{1.24}$$

The x_{α} should transform according to: $x_{\alpha}{}' = \frac{\partial x_{\alpha}{}'}{\partial x_{\beta}} x_{\beta}$ by the laws of partial differentiation-reminding the reader that the transformations are linear. We already know that $\frac{\partial x'_{\alpha}}{\partial x_{\beta}}$ is simply the matrix in 1.20 and we also know from 1.21 that this is the inverse of $\frac{\partial x'^{\alpha}}{\partial x^{\beta}}$. With this in mind we rewrite the transformations of the covariant and contravariant 4-vectors in terms of only the partial derivates with respect to the vector components with raised indices:

$$\begin{aligned} x'_{\alpha} &= \left(\frac{\partial x'_{\alpha}}{\partial x_{\beta}}\right) x_{\beta} = \left(\frac{\partial x^{\beta}}{\partial x^{\alpha\prime}}\right) x_{\beta} \text{ (covariant)} \\ x^{\alpha\prime} &= \left(\frac{\partial x^{\alpha\prime}}{\partial x^{\beta}}\right) x^{\beta} \text{ (contravariant)} \end{aligned} \tag{1.25}$$

Note that we can write these in terms of the Lorentz transform matrix Λ below:

$$x'_\alpha = \left(\frac{\partial x'_\alpha}{\partial x_\beta}\right) x_\beta = \left(\frac{\partial x^\beta}{\partial x^{\alpha\prime}}\right) x_\beta = (\Lambda^{-1})^\alpha_\beta x_\beta$$
$$x^{\alpha\prime} = \left(\frac{\partial x^{\alpha\prime}}{\partial x^\beta}\right) x^\beta = \Lambda^\alpha_\beta x^\beta \tag{1.26}$$

Note that any two arbitrary 4-vectors with these transformation properties:

$$A'_\mu = \frac{\partial x^\beta}{\partial x^{\mu\prime}} A_\beta \text{ and } B^{\mu\prime} = \frac{\partial x^{\mu\prime}}{\partial x^\alpha} B^\alpha$$

will form an invariant:

$$A'_\mu B^{\mu\prime} = \left(\frac{\partial x^\beta}{\partial x^{\mu\prime}}\right)\left(\frac{\partial x^{\mu\prime}}{\partial x^\alpha}\right) A_\beta B^\alpha$$

$$= \delta^\beta_\alpha A_\beta B^\alpha = A_\alpha B^\alpha$$

So we can form invariants from any two such vectors with the contravariant and covariant transformation properties $A_\mu B^\mu = A \cdot B$, from now we define a Lorentz scalar product using two 4-vectors A and B to be $A \cdot B = A_0 B^0 - \vec{A} \cdot \vec{B}$ where $\vec{A} \cdot \vec{B}$ is the spatial scalar product, this establishes the notation for the future.

1.5 Examples of 4-vectors: The 4-gradient

We will now identify other 4-vectors with these transformation properties and construct invariants from them as well; so far we have used only the co-ordinates of an event in space time. Let us begin with the operator $\frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial(ct)}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) = (\frac{\partial}{\partial(ct)}, \vec{\nabla})$ this is known as the 4-gradient . Using the rules of partial differentiation it is easy to see that:

$$\frac{\partial}{\partial x^{\mu\prime}} = \left(\frac{\partial x^\alpha}{\partial x'^\mu}\right)\frac{\partial}{\partial x^\alpha}$$

We know however from Eqn. 1.25 these are the properties of the *inverse* Lorentz transformation, applying to covariant vectors. Thus, $\frac{\partial}{\partial x^{\mu\prime}}$ which is a derivative with respect to a contravariant vector transforms like a *covariant* 4-vector, we therefore denote it by ∂_μ-the lowered index representing its' transformation properties. Let us now examine the derivative operator with respect to the components of the covariant four vector

$\frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial(ct)}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}) = \left(\frac{\partial}{\partial(ct)}, -\vec{\nabla}\right)$, by the rules of partial differentiation the transformed operator (in the primed frame) is:

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial}{\partial x_\alpha} = \left(\frac{\partial x^{\mu\prime}}{\partial x^\alpha}\right)\frac{\partial}{\partial x_\alpha}$$

where we have used the results in Eqn. 1.25 to infer replacing $\frac{\partial x_\alpha}{\partial x'_\mu}$ by $\frac{\partial x^{\mu\prime}}{\partial x^\alpha}$. Since this transforms like a contravariant vector and so we denote it by ∂^μ, with a raised index. We now write the Lorentz invariant operator

$$\partial_\mu \partial^\mu = \frac{\partial}{\partial x^{\mu\prime}} \cdot \frac{\partial}{\partial x'_\mu} = \frac{\partial}{\partial x^\mu} \cdot \frac{\partial}{\partial x_\mu} = \frac{\partial^2}{\partial(ct)^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$
$$= \frac{\partial^2}{\partial(ct)^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$

and is denoted, often by $\Box$, the d'Alembertian operator .

To summarize, $\partial'^\mu = \frac{\partial}{\partial x^{\mu\prime}} = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial}{\partial x_\alpha} = \left(\frac{\partial x^{\mu\prime}}{\partial x^\alpha}\right)\frac{\partial}{\partial x^\alpha}$, as an example the following equations represent a boost in the x direction for ∂^μ

$$\frac{\partial}{\partial x'_0} = \gamma\frac{\partial}{\partial x_0} - \gamma\beta\frac{\partial}{\partial x_1}$$
$$\frac{\partial}{\partial x'_1} = \gamma\frac{\partial}{\partial x_1} - \gamma\beta\frac{\partial}{\partial x_0}$$
$$\frac{\partial}{\partial x'_2} = \frac{\partial}{\partial x_2}$$
$$\frac{\partial}{\partial x'_3} = \frac{\partial}{\partial x_3}$$

and $\partial'{}_\mu = \frac{\partial}{\partial x^{\mu\prime}} = (\frac{\partial x^\alpha}{\partial x'^\mu})\frac{\partial}{\partial x^\alpha}$, once again a boost along x is given as an example

$$\frac{\partial}{\partial x^{0\prime}} = \gamma\frac{\partial}{\partial x^0} + \gamma\beta\frac{\partial}{\partial x^1}$$
$$\frac{\partial}{\partial x^{1\prime}} = \gamma\frac{\partial}{\partial x^1} + \gamma\beta\frac{\partial}{\partial x^0}$$
$$\frac{\partial}{\partial x^{2\prime}} = \frac{\partial}{\partial x^2}$$
$$\frac{\partial}{\partial x^{3\prime}} = \frac{\partial}{\partial x^3}$$

and finally the invariant d'Alembertian operator is easy to verify :

$$\partial_\mu \partial^\mu = \partial_\mu{}' \partial^{\mu\prime} = \Box = \frac{\partial^2}{\partial(ct)^2} - \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial(ct')^2} - \vec{\nabla}' \cdot \vec{\nabla}' \qquad (1.27)$$

1.6 Examples of 4-vectors: 4-momentum

In this section we will show how the energy and momentum of a particle in motion can be used to form a 4-vector . We consider a particle of rest mass m with velocity vector $\vec{v} = \frac{d\vec{x}}{dt}$ and construct the Lorentz-invariant:

$$\begin{aligned} ds &= \sqrt{c^2dt^2 - dx^2 - dy^2 - dz^2} = cdt\sqrt{1 - \frac{1}{c^2}\left(\frac{dx}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dy}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dz}{dt}\right)^2} \\ &= cdt\sqrt{1 - \frac{\vec{v}\cdot\vec{v}}{c^2}} \end{aligned}$$

We now denote $\frac{ds}{c} = dt\sqrt{1 - \frac{\vec{v}\cdot\vec{v}}{c^2}}$ by $d\tau$ the proper time or the time elapsed in the particle's own rest frame. Also note that $d\tau = \frac{dt}{\gamma}$ where $\gamma = \frac{1}{\sqrt{1 - \frac{\vec{v}\cdot\vec{v}}{c^2}}}$

Note that since $d\tau$ is an *invariant* quantity we note that

$$\begin{aligned} \frac{dx^0}{d\tau} &= \frac{dx_0}{dt}\frac{dt}{d\tau} = \gamma c \\ \frac{d\vec{x}}{d\tau} &= \frac{d\vec{x}}{dt}\frac{dt}{d\tau} = \gamma\vec{v} \end{aligned}$$

form the components of a contravariant vector, by virtue of the fact that $(x^0, \vec{x})$ forms a contravariant 4-vector and that therefore $(dx^0, d\vec{x})$ divided by an invariant results in $\left(\frac{dx^0}{d\tau}, \frac{d\vec{x}}{d\tau}\right)$ which must then also be a contravariant 4-vector. We call $\left(\frac{dx^0}{d\tau}, \frac{d\vec{x}}{d\tau}\right) = (\gamma c, \gamma\vec{v})$ the "4-velocity", denoted by v^α. We can form another contravariant vector from v^α by multiplying it by the particles mass this we denote by: $p^\alpha = mv^\alpha$:

$$p^\alpha = (m\gamma c, m\gamma\vec{v}) \tag{1.28}$$

Since we've already determined that p^α is a contravariant 4-vector, the following is an invariant:

$$p_\alpha p^\alpha = m^2(c^2 - v^2)\frac{1}{1 - \frac{v^2}{c^2}} = m^2c^2\left(1 - \frac{v^2}{c^2}\right)\frac{1}{\left(1 - \frac{v^2}{c^2}\right)} = m^2c^2 \tag{1.29}$$

Thus m^2c^2 is an invariant quantity. We note that $p^{\alpha'}$s spatial part is directly proportional to the components of the momentum . Let us examine the time component p^0:

$$\begin{aligned} m\gamma c &= mc\frac{1}{\left(1-\frac{v^2}{c^2}\right)^{1/2}} = mc(1+\frac{1}{2}\frac{v^2}{c^2}+\cdots) \\ &= mc + \frac{1}{2}m\frac{v^2}{c} + \cdots \end{aligned} \tag{1.30}$$

We identify the second term as the kinetic energy divided by c and the first term as the rest energy mc^2 divided by c. We note therefore that $p^\alpha = (\frac{E}{c}, \vec{p})$ where $\vec{p} = m\vec{u}\gamma$ and that $p^\alpha = (\frac{E}{c}, p)$ forms a contravariant 4-vector. In any process the total of each of the components of these 4-vectors–which are energy and momentum–are conserved, and can be used to form the invariant $p_\alpha p^\alpha$. Where only one particle is being considered this invariant is simply $\frac{E^2}{c^2} - p^2 = m^2c^2$. In the next equation we remind the reader of how the covariant 4-vector p_α can be formed from the action of the metric tensor on the contravariant 4-vector p^β and the different equivalent ways in which the individual components of the covariant may be expressed.

$$\begin{aligned} p_\alpha = g_{\alpha\beta}p^\beta = (\frac{E}{c}, -\vec{p}) = (\frac{E}{c}, -p_x, -p_y, -p_z) = (p_0, p_1, p_2, p_3) \\ = (p^0, -p^1, -p^2, -p^3) \end{aligned} \tag{1.31}$$

The invariants and conservation of 4-momenta are used in relativistic calculations and simplify the problems of viewing relativistic processes in different frames of reference greatly.

Finally, we make a note of another 4-vector related to the 4-momentum which occurs in quantum mechanics. Consider the free wave solution of the Schrödinger equation $\psi \propto e^{-i\frac{Et}{\hbar}+i\frac{\vec{p}\cdot\vec{x}}{\hbar}}$ which can also be written as $e^{-i\frac{\frac{E}{c}ct}{\hbar}+i\frac{\vec{p}\cdot\vec{x}}{\hbar}}$ where we now take E to be the total relativistic energy. The exponential factor is then $\frac{E}{\hbar} = \omega$ and $\frac{\vec{p}}{\hbar} = \vec{k}$ (the wave number) we can define the 4-vector

$$k^\mu = (\frac{\omega}{c}, \vec{k}) = (\frac{E}{\hbar}, \frac{\vec{p}}{\hbar}) \tag{1.32}$$

thus $\psi(x,t) = e^{-ik\cdot x} = e^{-ik_\mu x^\mu}$ with invariant $k_\mu x^\mu$. Not surprisingly this shows that the phase is a Lorentz-invariant quantity (once the total relativistic energy of the wave is used for E) i.e. no matter which frame you look at it in, the wave is at the same point of its evolution (it may be stretched or squished but a crest is a crest and a trough a trough in all frames of reference).

1.7 Tensors and their transformation properties

It is possible to form objects with more complicated transformation properties. Take for example every possible product of every possible component of the two contravariant vectors A and B, $A^\mu B^\nu$. Each component of the vectors transforms contravariantly. Consider now evaluating each of the components of A and B in a different Lorentz frame:

$$A^{\mu\prime} B^{\nu\prime} = \frac{\partial x^{\mu\prime}}{\partial x^\beta} A^\beta \frac{\partial x^{\nu\prime}}{\partial x^\alpha} B^\alpha = \left(\frac{\partial x^{\mu\prime}}{\partial x^\beta}\right) \left(\frac{\partial x^{\nu\prime}}{\partial x^\alpha}\right) A^\beta B^\alpha$$

The product $A^\beta B^\alpha$ is known as a tensor of rank 2 . We can write this as one object with two indices e.g. $T^{\mu\nu}$ with the transformation property:

$$T^{\mu\nu\prime} = \frac{\partial x^{\mu\prime}}{\partial x^\alpha} \frac{\partial x^{\nu\prime}}{\partial x^\beta} T^{\alpha\beta} \tag{1.33}$$

Note that because T and the partial derivatives carry two indices we can represent each by a matrix with the convention that the first(upper) index denotes a row and the second(lower) index represents a column for T. Since each index can take on 4 values (0, 1, 2, 3), these can be written as 4×4 matrices. Recall the way matrix multiplication is defined using arbitrary matrices A, B, and C: $A^{\tau\sigma} = B^{\tau\delta} C^{\delta\sigma}$ each $\tau\sigma$th element of the resulting matrix A is made by taking the the scalar product of the τth row of B with the σth column of the matrix C (as an example). The order of indices in 1.33 will thus tell us that if we denote the matrix of partial derivatives by M then the transformation in Eqn. 1.33 can be represented as matrix multiplication in the form:

$$T' = MTM^T \tag{1.34}$$

Where in the above equation the superscript ^{T} denotes the transpose of $\mathbb{M}$. Here is another example of a tensor of rank 2 :

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

where $F^{\mu\nu}$ is formed from partial derivatives of a vector A^α.

Another hypothetical example is $A^\mu{}_\nu = \partial^\mu B{}_\nu$ which is a mixed second rank tensor: one of the indices transforms covariantly and the other contravariantly. Note also the operation:

$$g_{\mu\nu} L^{\alpha\nu} = L^\alpha{}_\mu \tag{1.35}$$

will change a tensor with two contravariant indices to a tensor with one covariant and one contravariant index, and the following

$$g_{\mu\nu}g_{\beta\alpha}L^{\alpha\nu} = L_{\beta\mu} \tag{1.36}$$

lowers both indices of a second rank tensor. Finally the operation of summing over all indices is called a contraction, resulting in a Lorentz invariant.

$$\begin{aligned}
&T_{\mu\nu}T^{\mu\nu} = T_{00}T^{00} + T_{01}T^{01} + T_{02}T^{02} + T_{03}T^{03} + \\
&T_{10}T^{10} + T_{11}T^{11} + T_{12}T^{12} + T_{13}T^{13} + \\
&T_{20}T^{20} + T_{21}T^{21} + T_{22}T^{22} + T_{23}T^{23} + \\
&T_{30}T^{30} + T_{31}T^{31} + T_{32}T^{32} + T_{33}T^{33}
\end{aligned} \tag{1.37}$$

The reader can verify that this is actually true by inserting the partial derivatives that form the transformation matrices into the expression.

1.8 Infinitesimal Lorentz transformations of 4-vectors: definition of $\epsilon^{\mu}{}_{\nu}$

We now take a bit of a detour and define the generators of infinitesimal Lorentz transformations. These will be particularly useful in chapter 3, but an appropriate place for them is here, since this first chapter is a revision of selected topics in special relativity.

Recall the following definition for a Lorentz transformation for any 4-vector (here we use the p^{μ}:

$$p^{\nu\prime} = \Lambda^{\nu}{}_{\mu}p^{\mu} \tag{1.38}$$

where we have used the definition:

$$\Lambda^{\nu}{}_{\mu} = \frac{\partial x^{\nu\prime}}{\partial x^{\mu}} \tag{1.39}$$

Now consider

$$\Lambda^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \epsilon^{\nu}{}_{\mu} \tag{1.40}$$

which we claim represents an infinitesimal Lorentz transformation where $\epsilon^{\nu}{}_{\mu}$ is an infinitesimal matrix. Our 4-momentum then transforms like

$$p^{\nu\prime} = ({\delta^{\nu}}_{\mu} + {\epsilon^{\nu}}_{\mu})p^{\mu} \tag{1.41}$$

Using the parameterization in Eqns. 1.5-1.6 we paramaterize our Lorentz transformation as $\tanh(\Delta\omega) = \Delta\beta$ representing an infinitesimal velocity along the x-axis. We do this by defining ${\epsilon^{\nu}}_{\mu}(x)$ first

$${\epsilon^{\nu}}_{\mu} = \Delta\omega_x \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.42}$$

and then write down what we claim represents an infinitesimal Lorentz transformation explicitly using Eqn. 1.42

$${\delta^{\nu}}_{\mu} + {\epsilon^{\nu}}_{\mu} = \begin{pmatrix} 1 & -\Delta\omega_x & 0 & 0 \\ -\Delta\omega_x & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.43}$$

We can now write the infinitesimal transform of the 4-momentum explicitly:

$$\begin{aligned} p^{\nu\prime} &= ({\delta^{\nu}}_{\mu} + {\epsilon^{\nu}}_{\mu})p^{\mu} \\ p^{0\prime} &= p^0 - \Delta\omega_x p^1 \\ p^{1\prime} &= p^1 - \Delta\omega_x p^0 \\ p^{2\prime} &= p^2 \\ p^{3\prime} &= p^3 \end{aligned} \tag{1.44}$$

We can compare these $p^{0\prime}$ and $p^{1\prime}$ to a finite Lorentz transformation

$$p^{0\prime} = \gamma(p^0 - \beta p^1)$$

$$p^{1\prime} = \gamma(p^1 - \beta p^0)$$

if β is tiny it is easy to see how Eqn. 1.44 follows. It should be clear that $\gamma \approx 1$ in this limit. It is conventional to write ${\epsilon^{\nu}}_{\mu}$ as $\Delta\omega {I^{\nu}}_{\mu}$ where ${I^{\nu}}_{\mu}$ has -1 and 0s as its' entries, the full set of matrices for boosts along x, y and z then follow:

$$\epsilon^{\nu}{}_{\mu}(x) = \Delta\omega_x \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \Delta\omega_x I^{\nu}{}_{\mu}(x) \tag{1.45}$$

$$\epsilon^{\nu}{}_{\mu}(y) = \Delta\omega_y \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \Delta\omega_y I^{\nu}{}_{\mu}(y) \tag{1.46}$$

$$\epsilon^{\nu}{}_{\mu}(z) = \Delta\omega_z \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} = \Delta\omega_z I^{\nu}{}_{\mu}(z) \tag{1.47}$$

In Eqns. 1.45-1.47 we have defined matrices for infinitesimal boosts and the associated $I^{\nu}{}_{\mu}(x_i)$ matrices (i =1,2,3 for x, y, z). The equation 1.41 for an infinitesimal Lorentz transformation can be written as:

$$p^{\nu\prime} = (\delta^{\nu}{}_{\mu} + \Delta\omega I^{\nu}{}_{\mu})p^{\mu} \tag{1.48}$$

Note how the application of the metric tensor raises and lowers indices. We do this for $\epsilon(x)$ as an example. The reader is first reminded that the lower index in $\epsilon^{\nu}{}_{\mu}$ is the second index. Thus $\epsilon_{\alpha\mu} = g_{\alpha\nu}\epsilon^{\nu}{}_{\mu}$, more specifically lets pick the 01^{th} element: $\epsilon_{01} = g_{0\nu}\epsilon^{\nu}_{1} = g_{00}\epsilon^{0}_{1} + g_{01}\epsilon^{1}{}_{0} + \cdots$ this operation is easily seen to yield:

$$g_{\alpha\nu}\epsilon^{\nu}{}_{\mu} = \epsilon_{\alpha\mu}(x) = \Delta\omega_x \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.49}$$

One must be a little careful when raising the second index (μ) of $\epsilon^{\nu}{}_{\mu}$ to obtain $\epsilon^{\nu\beta}$ for now we have to sum over the *second* index of $\epsilon^{\nu}{}_{\mu}$. Thus $\epsilon^{\nu\beta} = g^{\mu\beta}\epsilon^{\nu}{}_{\mu}$, explicitly we can do this for the 01^{th} element again : $\epsilon^{01} = g^{\mu 1}\epsilon^{1}{}_{\mu} = g^{11}\epsilon^{1}{}_{0}$, where only the non-zero terms have been shown.

$$g^{\beta\mu}\epsilon^{\nu}{}_{\mu} = \epsilon^{\nu\beta}(x) = \Delta\omega_x \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.50}$$

It is easy to see that *all* the $\epsilon_{\mu\nu}$ (and $\epsilon^{\mu\nu}$) are anti-symmetric with respect to the interchange of the indices μ and ν but not the ϵ^{ν}_{μ}). We simply write :

$$\begin{aligned} \epsilon_{\mu\nu} &= -\epsilon_{\nu\mu} \\ \epsilon^{\mu\nu} &= -\epsilon^{\nu\mu} \end{aligned} \tag{1.51}$$

So far we have simply defined what we claim to represent infinitesimal boosts (Lorentz transformations) for 4-vectors and have defined the properties of the matrices used. We will now prove by successive applications of infinitesimal boosts (with $\Delta\beta = \Delta\omega$) that we recover the usual Lorentz transformation equations for a 4-vector (see Eqns 1.1,1.5, 1.6). The following expression represents an infinite number of infinitesimal boosts where $\Delta\omega = \frac{\omega}{N}$ as $N \to \infty$:

$$\lim_{N\to\infty} (\delta^{\nu}{}_{\mu} + \frac{\omega}{N} I^{\nu}{}_{\mu})^N \tag{1.52}$$

The Eqn. 1.52 is simply Eqn. 1.52 with $\Delta\omega$ replaced with $\frac{\omega}{N}$ where the limit of large N is understood. Note that the expression 1.52 is simply the definition of the exponential function:

$$e^{\omega I} = \lim_{N\to\infty} (\delta^{\nu}{}_{\mu} + \frac{\omega}{N} I^{\nu}{}_{\mu})^N \tag{1.53}$$

We will now validate a boost along x for which we use the appropriate matrix :

$$I^{\nu}{}_{\mu}(x) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = I$$

Note that trivially $I^2(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

and consequently $I^n(x)$ for n odd $= \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Thus the exponential in Eqn. 1.53 is easy to expand and we see that it is represented by the following series:

$$e^{\omega I} = \mathbb{I} + \omega I + \frac{\omega^2 I^2}{2!} + \frac{\omega^3 I^3}{3!} + \cdots$$

$$= \begin{pmatrix} 1 + \frac{\omega^2}{2!} + \frac{\omega^4}{4} + \cdots & -\omega - \frac{\omega^3}{3!} - \frac{\omega^5}{5!} - \cdots & 0 & 0 \\ -\omega - \frac{\omega^3}{3!} - \frac{\omega^5}{5!} - \cdots & 1 + \frac{\omega^2}{2} + \frac{\omega^4}{4} + \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Collecting all the entries in the matrix we can identify Taylor expansions of the hyperbolic functions $\sinh \omega$ and $\cosh \omega$; this expanded matrix is simply

$$\begin{pmatrix} \cosh \omega & -\sinh \omega & 0 & 0 \\ -\sinh \omega & \cosh \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and with $\tanh(\omega) = \beta$ this simply defines Lorentz boost along x as in Eqns. 1.5-1.6.

1.9 Useful notation for calculating vector identities

Prior to looking at the transformations properties of current and charge densities, electromagnetic fields and potentials the calculation of some vector identities is presented. A notation using the Levi-Cevita and Kronecker symbols is used providing, it is hoped, a convenient shorthand for the student for understanding the calculation presented here and in other texts.

To introduce this new notation we look at the curl of a vector field. Written in the "usual way" this is $\vec{\nabla} \times \vec{C}$. We now write the i^{th} component of the curl in the following fashion remembering that repeated indices are summed over and free indices take on all possible values in the relevant context:

$$\epsilon_{ijk} \partial_j C_k \tag{1.54}$$

where ϵ_{ijk} is known as the Levi-Cevita tensor density and its properties are as follows:

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any index is repeated} \\ 1 & \text{for } \textit{even} \text{ permutations of the indices i.e. } \epsilon_{123} = \epsilon_{312} = \epsilon_{231} = 1 \\ -1 & \text{for } \textit{odd} \text{ permutations of the indices i.e. } \epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \end{cases}$$

where summation over repeated indices is implied. As an example setting $i = 1$ we have

$$\epsilon_{1jk}\partial_j C_k = \epsilon_{123}\partial_2 C_3 + \epsilon_{132}\partial_3 C_2 = \partial_y C_z - \partial_z C_y$$

where $\partial_z = \frac{\partial}{\partial z}$ etc.

The divergence of a vector field $\vec{\nabla}\cdot\vec{C}$ can now be expressed quite simply using the repeated index summation notation:

$$\partial_i C_i = \vec{\nabla}\cdot\vec{C} = \partial_1 C_1 + \partial_2 C_2 + \partial_3 C_3 = \partial_x C_x + \partial_y C_y + \partial_z C_z$$

The i^{th} component of the divergence of the curl of a vector field $\vec{C}$ $\vec{\nabla}\cdot(\vec{\nabla}\times\vec{C})$ can be written using this notation as well, it should be clear then that this is:

$$\partial_i \epsilon_{ijk}\partial_j C_k \tag{1.55}$$

Rearranging Eqn. 1.55 we obtain $\epsilon_{ijk}\partial_i\partial_j C_k$ we note that: $\partial_i\partial_j C_k = \partial_j\partial_i C_k$ (exchanging the order of partial derivatives) it is easy to see then that any component of $\vec{C}$ will occur *twice* in the summation; once with an even permutation of the indices of ϵ_{ijk} and once with an odd permutation and hence with opposite signs and so we get zero for the divergence of the curl of a vector field.

$$\vec{\nabla}\cdot(\vec{\nabla}\times\vec{C}) = 0 \tag{1.56}$$

The next operation we consider is: $\vec{\nabla}\times(\vec{\nabla}\times\vec{C})$. It is easy to see that the l^{th} component of the result is given by:

$$\epsilon_{lmi}\epsilon_{ijk}\partial_m\partial_j C_k \tag{1.57}$$

It is easy to verify (by plugging in indices explicitly) that:

$$\epsilon_{lmi}\epsilon_{ijk} = \epsilon_{ilm}\epsilon_{ijk} = \delta_{lj}\delta_{mk} - \delta_{mj}\delta_{lk}$$

. So we have:

$$(\delta_{lj}\delta_{mk} - \delta_{mj}\delta_{lk})\partial_m\partial_j C_k \tag{1.58}$$

The cases for which the term in the brackets in Eqn. 1.58 is not zero are $l = j$, $m = k$ with $m \neq j$, $l \neq k$ and then $m = j$, $l = k$, with $l \neq j$ $m \neq k$. The first case gives us:

$$\delta_{lj}\delta_{mk}\partial_m\partial_j C_k = \partial_k\partial_j C_k = \partial_j\partial_k C_k$$

these are simply the components of $\vec{\nabla}(\vec{\nabla} \cdot \vec{C})$. The second case gives us:

$$-\delta_{mj}\delta_{lk}\partial_m\partial_j C_k = -\partial_j\partial_j C_k$$

this clearly represents ∇^2 operating on *each* component of $\vec{C}$. Combining the two cases we obtain the well known result:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{C} = -\nabla^2\vec{C} + \vec{\nabla}(\vec{\nabla} \cdot \vec{C}) \tag{1.59}$$

Finally we consider the curl of the gradient of a scalar field , $\nabla \times \nabla\phi$. The i^{th} component (with i taking all values 1,2, and 3) is $\epsilon_{ijk}\partial_j\partial_k\phi$ it it is straight forward to see that because of the interchangeability of the order of partial derivatives each term appears twice with a different sign due to the order of j and k and this will sum to zero. Thus $\nabla \times \nabla\phi = 0$ for any scalar field ϕ. The three results

$$\begin{aligned}
&\vec{\nabla} \cdot \vec{\nabla} \times \vec{C} = 0 \\
&\vec{\nabla} \times \vec{\nabla} \times \vec{C} = -\nabla^2\vec{C} + \vec{\nabla}(\vec{\nabla} \cdot \vec{C}) \\
&\vec{\nabla} \times \vec{\nabla}\phi = 0
\end{aligned} \tag{1.60}$$

will be used in the following sections. It is hoped that the notation in this section helps the student with this and other texts using vector identities extensively.

1.10 Examples of 4-vectors: the 4-current

In this section we will show that the current density $\vec{J}$ and the charge density ρ can be combined to form the components of a 4-vector. We begin with two of Maxwell's equations :

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho \tag{1.61}$$

where ρ is the charge density (charge per volume) and

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi \vec{J}}{c} \tag{1.62}$$

Taking the divergence of Eqn. 1.62

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) = \frac{1}{c}\frac{\partial(\vec{\nabla} \cdot \vec{E})}{\partial t} + \frac{4\pi}{c}(\vec{\nabla} \cdot \vec{J})$$

Recall that the divergence of a curl equals zero (Eqn. 1.60)so the left hand side is zero and we have:

$$\frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{E}) + \frac{4\pi}{c}\vec{\nabla} \cdot \vec{J} = 0$$

But we know that $\vec{\nabla} \cdot \vec{E} = 4\pi\rho$ therefore:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \tag{1.63}$$

This is the continuity equation expressing charge conservation. We note that the continuity equation must hold in *all* frames of reference, if the derivatives and the current and charge densities are Lorentz transformed (appropriately) the same relation should hold, so the right hand side (0) is a Lorentz invariant. We can "spot" the components of one 4-vector in the Eqn. 1.63 that is of the covariant $\partial_\alpha = (\frac{\partial}{\partial ct}, \vec{\nabla})$ which leads us to conclude that $(\rho c, \vec{J})$ are the components of a contravariant vector J^α. Consequently we know that $\partial_\alpha J^\alpha$ is an invariant and equal to zero in all frames.

Another way to think about this is to consider the 4-vector formed by multiplying the 4-velocity by ρ_0. This gives us $(\rho_0 \gamma c, \rho_0 \gamma \vec{v})$. We know that this transforms contravariantly by comparing to 1.28. We also know that a charge density times a velocity is exactly what $\vec{J}$ is for a point like particle. Once again this leads us to the conclusion that $J^\alpha = (c\rho, \vec{J})$ is a contravariant vector and we heretofore refer to J^α as the 4-current .

1.11 Combining the magnetic vector potential $\vec{A}$ and the scalar potential ϕ into a 4-vector

We will now demonstrate that the electrostatic scalar potential $\phi(x, t)$ and the vector potential $\vec{A}(x, t)$ can also be combined to form the components of a four vector. To do this we begin by considering the following two of Maxwell's equations :

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{J} \tag{1.64}$$

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \tag{1.65}$$

By expressing the $\vec{B}$ field as the curl of the magnetic vector potential we can rewrite Eqn. 1.65 in the following form:

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} = -\frac{1}{c}\vec{\nabla} \times \left(\frac{\partial \vec{A}}{\partial t}\right) \tag{1.66}$$

We can remove the $\vec{\nabla}\times$ from the left and right hand sides if we promise to add the gradient of a scalar field $= -\vec{\nabla}\phi$ to the right hand side since the relation $\vec{\nabla} \times \vec{\nabla}\phi = 0$ holds for any arbitrary scalar field ϕ. We will next identify this to be the scalar potential ϕ recalling that $\vec{E} = -\vec{\nabla}\phi$ we obtain:

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi \tag{1.67}$$

Next we use $\vec{B} = \vec{\nabla}\times\vec{A}$, and substitute for $\vec{E}$ from Eqn. 1.67 in Eqn. 1.64 to obtain:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \frac{1}{c}\frac{\partial}{\partial t}\left(-\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi\right) + \frac{4\pi}{c}\vec{J} \tag{1.68}$$

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = -\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t} - \frac{1}{c}\vec{\nabla}\left(\frac{\partial \phi}{\partial t}\right) + \frac{4\pi}{c}\vec{J} \tag{1.69}$$

Using $\vec{\nabla} \times \vec{\nabla} \times \vec{A} = -\nabla^2\vec{A} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A})$ and rearranging :

$$-\nabla^2\vec{A} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) = -\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} - \frac{1}{c}\vec{\nabla}\left(\frac{\partial \phi}{\partial t}\right) + \frac{4\pi}{c}\vec{J}$$

Rearranging further:

$$-\nabla^2\vec{A} + \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} + \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}\right) = \frac{4\pi}{c}\vec{J}$$

Now recall Eqn. 1.67 :

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi$$

and take the divergence of of the left and right hand side

$$\vec{\nabla} \cdot \vec{E} = -\frac{1}{c}\frac{\partial(\vec{\nabla} \cdot \vec{A})}{\partial t} - \vec{\nabla}^2\phi$$

$$-\frac{1}{c}\frac{\partial(\vec{\nabla} \cdot \vec{A})}{\partial t} - \nabla^2\phi = 4\pi\rho$$

where we use the fact that the divergence of the electric field is 4π times the charge density. We now recognize that we have a pair of coupled differential equations for $\vec{A}$ and ϕ:

$$-\nabla^2 A + \frac{1}{c^2}\frac{\partial^2 A}{\partial t^2} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}) = \frac{4\pi \vec{J}}{c} \tag{1.70}$$

$$-\nabla^2\phi - \frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{A}) = 4\pi\rho \tag{1.71}$$

First let us recall that the right hand sides of Eqns. 1.71 and 1.70 have the time and space components of the 4-vector J^α multiplied by $\frac{4\pi}{c}$. Note that the $-\nabla^2 A + \frac{1}{c^2}\frac{\partial^2 A}{\partial t^2}$ is simply $\Box\vec{A}$, as we know the $\Box$ is a Lorentz invariant operator. If $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}$ were $= 0$, then $\vec{\nabla} \cdot \vec{A} = -\frac{1}{c}\frac{\partial \phi}{\partial t}$ and we could use this in Eqn. 1.71 and obtain:

$$-\nabla^2\phi + \frac{1}{c}\frac{\partial^2 \phi}{\partial t^2} = 4\pi\rho \tag{1.72}$$

This is just $\Box\phi = 4\pi\rho$. We would then have the two equations:

$$\begin{aligned} \Box\vec{A} &= \frac{4\pi\vec{J}}{c} \\ \Box\phi &= 4\pi\rho \end{aligned} \tag{1.73}$$

Since we would have the components of a 4-vector on the right hand side and an invariant operator acting on $\vec{A}$ and ϕ on the left this would indicate that $\vec{A}$ and ϕ could be combined into a 4-vector. So what should we do with $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0$? As we shall it can be made to vanish without changing any physical observables namely $\vec{B}$ and $\vec{E}$.

We begin with $\vec{B} = \vec{\nabla} \times \vec{A}$ and recall that for any scalar field $\mathcal{G}(\vec{x}, t)$ the gradient of its curl is zero: $\vec{\nabla} \times (\vec{\nabla}\mathcal{G}(\vec{x}, t)) = 0$. This means that we can always change $\vec{A}$ to $\vec{A} + \vec{\nabla}\mathcal{G}(\vec{x}, t)$ where $\mathcal{G}$ is an arbitrary scalar function of $\vec{x}$ and t, since this will not change the physically measurable $\vec{B}$ field. But from Eqn 1.67 we know that any change in $\vec{A}$ will also change the physically measureable $\vec{E}$:

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla}\mathcal{G}) - \vec{\nabla}\phi = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}(\phi + \frac{1}{c}\frac{\partial \mathcal{G}}{\partial t}) \tag{1.74}$$

Thus the change in the vector potential has given rise to an extra term $-\vec{\nabla}\frac{1}{c}\frac{\partial \mathcal{G}}{\partial t}$, so although the $\vec{B}$ field remained unchanged we have changed the electric field. However if when changing $\vec{A}$ to $\vec{A} + \vec{\nabla}\mathcal{G}(\vec{x}, t)$, we also change $\phi \to \phi - \frac{1}{c}\frac{\partial \mathcal{G}}{\partial t}$ then the extra term of $-\vec{\nabla}\frac{1}{c}\frac{\partial \mathcal{G}}{\partial t}$ will cancel out leaving the physically measureable $\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi$ unchanged.

Recall that the term we want to make disappear (discussion between Eqn. 1.70 and Eqn. 1.73) is $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}$. With the modification to $\vec{A}$ and ϕ this term changes from $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}$ to :

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} + \nabla^2\mathcal{G}(\vec{x}, t) - \frac{1}{c^2}\frac{\partial^2 \mathcal{G}(\vec{x}, t)}{\partial t^2} \tag{1.75}$$

Note that both $\nabla^2\mathcal{G}(\vec{x}, t) - \frac{1}{c^2}\frac{\partial^2 \mathcal{G}(\vec{x}, t)}{\partial t^2}$ and $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}$ are simply scalar functions of $\vec{x}$ and t. Since $\mathcal{G}(x, t)$ is a completely arbitrary scalar field it is always possible to pick it such that $\nabla^2\mathcal{G}(\vec{x}, t) - \frac{1}{c^2}\frac{\partial^2 \mathcal{G}(\vec{x}, t)}{\partial t^2}$ cancels the term

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}$$

exactly. From this point onward we will always assume that this has been done and that consequently $\vec{\nabla} \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0$, this is known as the Lorenz gauge condition (not Lorentz but Lorenz, these were two different people!). This means that Eqn. 1.70 and Eqn. 1.71 will simply reduce to Eqn. 1.73, noting again that on the left side of Eqn. 1.73 the $\Box$ operator is Lorentz invariant and the right hand side has ρ and $\vec{J}$ which we already proved in Sec. 1.10 are components of a 4-vector. We can now summarize our knowledge of and conditions on $\vec{A}$ and ϕ:

1. A and ϕ form the components of a contravariant vector, we will refer to this as the 4-vector potential of electromagnetism; with $A^\mu = (\phi, \vec{A})$

2. The components of A^μ satisfy the equation:

$$\begin{aligned} \Box A^\mu &= +\frac{4\pi}{c} J^\mu \\ \text{with } J^\mu &= (\rho c, \vec{J}) \end{aligned} \tag{1.76}$$

3. Neccesary for items 1 and 2 (above) the Lorenz Gauge condition $\vec{\nabla}\cdot\vec{A}+\frac{1}{c}\frac{\partial\phi}{\partial t} = 0$ on the components of $A^\mu = (\phi, \vec{A})$ must hold. This is conveniently expressed in relativistically invariant form:

$$\partial_\mu A^\mu = 0 \tag{1.77}$$

1.12 Lorentz transformation of electromagnetic fields: $\vec{E}$ and $\vec{B}$ as components of a second rank tensor

It is natural given our experience of the previous sections to ask if the magnetic and electric fields could also be considered to be somehow components of a single object.

$$\begin{aligned} \vec{E} &= -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi \\ \vec{B} &= \vec{\nabla}\times\vec{A} \end{aligned} \tag{1.78}$$

We begin by noting that the right hand side of Eqn. 1.78 involve space and time derivatives of components of the components of the 4-vector A^α and thus we know that $\vec{E}$ and $\vec{B}$ cannot be the components of a 4-vector but must be derived from some object with more complicated transformation properties and remind the reader of the notation adopted for our 4-gradient operator:

$$\partial^\alpha = (\partial^0, -\vec{\nabla}) = (\partial^0, \partial^i)$$

where i denote spatial indices–Greek indices typically denote spatial and time components. We next write down the x-components of the E and B fields in Eqn. 1.78:

$$\begin{aligned} E_x &= -\frac{1}{c}\frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x} &&= -\partial^0 A^1 + \partial^1 A^0 \\ B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} &&= -\partial^2 A^3 + \partial^3 A^2 \end{aligned} \tag{1.79}$$

In the light of the above consider the second rank tensor (see discussion in 1.33) $F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha$. It is easily seen that elements $F^{10} = -F^{01} =$

E_x and $F^{23} = -F^{32} = -B_x$ as we calculated in Eqns. 1.79. It can easily be verified that this can be represented explicitly in matrix form as:

$$\partial^\alpha A^\beta - \partial^\beta A^\alpha = F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{1.80}$$

This is known as the field strength tensor of electromagnetism, it is anti-symmetric in the indices α and β and note that because of the transformation properties of ∂^μ and A^μ it is a tensor of rank 2 as defined in Sec. 1.33. With this definition and recalling that $J^\alpha = (c\rho, \vec{J})$ it is easy to verify that two of Maxwell's equations

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= 4\pi\rho \\ \vec{\nabla} \times \vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} &= \frac{4\pi \vec{J}}{c} \end{aligned} \tag{1.81}$$

can be summarized succintly in the next expression:

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} J^\beta \tag{1.82}$$

Note that we can transform Lorentz transform $F^{\mu\nu}$ according to the rules for a tensor, for simplicities sake we use a boost along x:

$$F^{\mu\nu\prime} = \frac{\partial x^{\mu\prime}}{\partial x^\beta} \frac{\partial x^{\nu\prime}}{\partial x^\alpha} F^{\beta\alpha}$$

As we discussed in Sec. 1.33 this is

$$\begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Multiplying this out, we have:

$$\begin{pmatrix} -\gamma\beta E_x & -\gamma E_x & -\gamma E_y + \gamma\beta B_z & -\gamma E_z - \gamma\beta B_y \\ \gamma E_x & \gamma\beta E_x & \gamma\beta E_y - \gamma B_z & \gamma\beta E_z + \gamma B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} -\gamma^2\beta E_x + \gamma^2\beta E_x & -\gamma^2\beta^2 E_x - \gamma^2 E_x & -\gamma E_y + \gamma\beta B_z & -\gamma E_z - \gamma\beta B_y \\ \gamma^2 E_x - \gamma^2\beta^2 E_x & -\gamma^2\beta E_x + \gamma^2\beta^2 E_x & \gamma\beta E_y - \gamma B_z & \gamma\beta E_z + \gamma B_y \\ \gamma E_y - \gamma\beta B_z & -\gamma\beta E_y + \gamma B_z & 0 & -B_x \\ \gamma E_z + \gamma\beta B_y & -\gamma\beta E_z - \gamma B_y & B_x & 0 \end{pmatrix}$$

Note that $\gamma^2\beta^2 - \gamma^2 = -1$ and $\gamma^2(1-\beta^2) = 1$. Using this we find that the transformed tensor in matrix form is equal to

$$\begin{pmatrix} 0 & -E_x & -(\gamma E_y - \gamma\beta B_z) & -(\gamma E_z + \gamma\beta B_y) \\ E_x & 0 & -(+\gamma B_z - \gamma\beta E_y) & (\gamma B_y + \gamma\beta E_z) \\ (\gamma E_y - \gamma\beta B_z) & (\gamma B_z - \gamma\beta E_y) & 0 & -B_x \\ (\gamma E_z + \gamma\beta B_y) & -(\gamma B_y + \gamma\beta E_z) & B_x & 0 \end{pmatrix} \tag{1.83}$$

Notice how E_x and B_x are unchanged but the fields perpendicular to the direction of motion have changed. Compare this with the un-transformed tensor in Eqn. 1.80.

We can also see how to constuct invariants from this. Consider $F_{\alpha\beta} = g_{\alpha\gamma}g_{\beta\sigma}F^{\gamma\sigma}$. Recall from Eqn. 1.36 that this requires the following matrix operation:

$$F_{\alpha\beta} = g_{\alpha\gamma}g_{\beta\sigma}F^{\gamma\sigma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

(Note: the third matrix is the transpose of $g_{\mu\nu}$ which is the same as $g_{\mu\nu}$). We know from the discussion toward the end of Sec. 1.33 that the contraction $F_{\alpha\beta}F^{\alpha\beta}$ is Lorentz invariant, so it is trivial to see that $= F_{\alpha\beta}F^{\alpha\beta} = F_{00}F^{00} + F_{01}F^{01} + ...F_{33}F^{33}$ is simply equal to $2(- \mid \vec{E} \mid^2 + \mid \vec{B} \mid^2)$ which is Lorentz invariant. The reader is encouraged to check this using the transformed electromagnetic tensor in Eqn. 1.83 and untransformed one 1.80.

Finally recall the form of Eqn. 1.82; we can also express the remaining two of Maxwell's equations in this spirit but first we have to go through some definitions. We begin by defining $\epsilon_{\gamma\theta\alpha\beta}$ (4-dimensional analogue of ϵ_{ijk}) with the following properties (same as $\epsilon^{\gamma\theta\alpha\beta}$)

$$\epsilon^{\gamma\theta\alpha\beta} = \epsilon_{\gamma\theta\alpha\beta} = \begin{cases} 0 & \text{if any of the indices are repeated} \\ 1 & \text{for all } \textit{even} \text{ permutations of } \alpha,\, \beta,\, \gamma,\, \theta, \\ -1 & \text{for all } \textit{odd} \text{ permutations of } \alpha,\, \beta,\, \gamma,\, \theta \end{cases}$$

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2}\epsilon^{\alpha\beta\gamma\theta}F_{\gamma\theta} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix} \tag{1.84}$$

Let us now see how to generate (this also serves as a cross check) some of these entries. Consider for example $1, 2^{th}$ component of $\mathcal{F}^{\alpha\beta}$, i.e. $\mathcal{F}^{12} = \frac{1}{2}\epsilon^{12\gamma\theta}F_{\gamma\theta}$. If either γ or $\theta = 1$ or 2 we'll get 0, so γ and θ have to be 0, 3. Using this information it is easy to see that

$$\mathcal{F}^{12} = \frac{1}{2}\epsilon^{1203}F_{03} + \frac{1}{2}\epsilon^{1230}F_{30} = \frac{1}{2}(\epsilon^{1203}(-E_z) + \epsilon^{1230}E_z)$$

which means that

$$\mathcal{F}^{12} = \frac{1}{2}((-1)(-1)E_z + (+1)E_z) = E_z$$

1 .

It is easy to see that the following two of Maxwell's equations

$$\vec{\nabla}\cdot\vec{B} = 0 \text{ and } \vec{\nabla}\times\vec{E} - \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0$$

can be compactly summarized in the following way: $\partial_\alpha\mathcal{F}^{\alpha\beta} = 0$. To test this claim we can explicitly do the sum over α for any given value of β, for example $\beta = 1$:

$$\partial_\alpha\mathcal{F}^{\alpha 1} = \partial_0\mathcal{F}^{01} + \partial_1\mathcal{F}^{11} + \partial_2\mathcal{F}^{21} + \partial_3\mathcal{F}^{31}$$

Switching explicitly to the co-ordinates (time and space) representing the indices we obtain:

$$\partial_0(-B_x) + \partial_y E_z - \partial_z E_y - \frac{1}{c}\frac{\partial B_x}{\partial t} + (\vec{\nabla}\times\vec{E})_x = 0$$

2 . This is clearly the x^{th} component of one of the Maxwell equations namely
3 $\vec{\nabla}\times\vec{E} - \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0$. So we can see easily that

$$\begin{aligned} &\vec{\nabla}\cdot\vec{B} = 0 \\ &\vec{\nabla}\times\vec{E} - \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0 \end{aligned} \tag{1.85}$$

4 can be summarized as :

$$\partial_\alpha\mathcal{F}^{\alpha\beta} = 0 \tag{1.86}$$

Note that equations 1.86 and 1.82 are referred to as Maxwell's equations in covariant form, convariant used in this context means that the equations retain their form (using the transformed tensors and vectors) in a different frame of reference.

It has already been shown that $F_{\mu\nu}F^{\mu\nu} = 2(\mid \vec{B} \mid^2 - \mid \vec{E} \mid^2)$ is a Lorentz invariant, it should be clear that $F_{\mu\nu}\mathcal{F}^{\mu\nu} = -4\vec{E}\cdot\vec{B}$ is also a Lorentz invariant. The tensor $\mathcal{F}^{\mu\nu}$ is known as the dual electromagnetic tensor [1]. Finally the reader is asked to bear in mind that matrices have been used to represent Lorentz transformations for convenience and ease of manipulation, and 4-vectors have been represented as column vectors for the same reason. We will encounter in the next chapter matrices of a different nature which operate on elements of a wave function that is a column vector of 4 entries. These will not form 4-vectors, they belong to an object of different nature with different rules for Lorentz transformations. Several excellent texts contain a detailed treatment of this topic [1 5].

1.13 Charged particle in an electromagnetic field

The goal of this section is to show how the motion of a charged particle in an electromagnetic field can be incorporated into a Hamiltonian, this will then be used in quantum mechanical Hamiltonians in later chapters. The claim is that the modified Hamiltonian (simply replace $\vec{p}$ by $\vec{p} - \frac{q\vec{A}}{c}$)

$$H = \frac{1}{2m}\left(\vec{p} - \frac{q\vec{A}}{c}\right)^2 + q\phi \tag{1.87}$$

where ϕ is the scalar potential $\phi = \phi(\vec{x}, t)$ and $\vec{A} = A(\vec{x}, t)$ is the vector potential (see Eqn. 1.78) contains the appropriate modification that will lead to the expression for the motion of a charged particle in an electromagnetic field i.e. :

$$\vec{F} = q\frac{\vec{v}}{c} \times \vec{B} + q\vec{E} \tag{1.88}$$

Here the reader is reminded that $\vec{F}$ is the rate of change of the mechanical momentum $m\dot{\vec{x}}$. Expanding Eqn 1.87 we obtain:

$$H = \frac{1}{2m}\vec{p}\cdot\vec{p} - q\frac{\vec{p}\cdot\vec{A}}{mc} + \frac{q^2}{2mc^2}\vec{A}\cdot\vec{A} + q\phi$$

We will now use Hamilton's equations [6] $\frac{\partial H}{\partial p} = \frac{d\vec{x}}{dt}$ and $\frac{\partial H}{\partial x} = -\frac{d\vec{p}}{dt}$ to show that this is true. Using the ith component of the first equation we obtain trivially:

$$\frac{dx_i}{dt} = \frac{p_i}{m} - \frac{qA_i}{cm} \tag{1.89}$$

Using the second of Hamilton's equations we obtain:

$$(\dot{\vec{p}})_i = (= -\frac{\partial H}{\partial x_i} = -\frac{\partial}{\partial x_i}\left(\frac{1}{2m}\vec{p}\cdot\vec{p} - \frac{q\vec{p}}{mc}\cdot\vec{A} + \frac{q^2\vec{A}\cdot\vec{A}}{2mc^2} + q\phi\right)$$

and in terms of the i^{th} component of the equation we have:

$$\dot{p_i} = -\frac{\partial H}{\partial x_i} = -\frac{\partial}{\partial x_i}(\frac{1}{2m}p_j p_j - \frac{q}{mc}p_j A_j + \frac{q^2}{2mc^2}A_j A_j + q\phi)$$

$$\dot{p_i} = -\frac{\partial H}{\partial x_i} = \frac{q}{mc}p_j\frac{\partial A_j}{\partial x_i} - \frac{q^2}{mc^2}A_j\frac{\partial A_j}{\partial x_i} - q\frac{\partial\phi}{\partial x_i} \tag{1.90}$$

The reader is reminded that in the development above we are dealing with the *canonical* momentum p. Now note from Eqn. 1.89, we have:

$$\frac{p_i}{m} - \frac{qA_i}{mc} = \frac{dx_i}{dt}$$

and so

$$\frac{p_i}{m} - \frac{dx_i}{dt} = \frac{qA_i}{mc}$$

using this we can substitute for $\frac{q}{mc}A_j$ (using index i in place of j) in Eqn. 1.90 to obtain

$$\dot{p_i} = \frac{q}{mc}p_j\frac{\partial A_j}{\partial x_i} - \frac{q}{c}\frac{q}{c}\frac{A_j}{m}\frac{\partial A_j}{\partial x_i} - q\frac{\partial\phi}{\partial x_i}$$

which is

$$= \frac{q}{mc}p_j\frac{\partial A_j}{\partial x_i} - \frac{q}{c}\left(\frac{p_j}{m} - \frac{dx_j}{dt}\right)\frac{\partial A_j}{\partial x_i} - q\frac{\partial\phi}{\partial x_i}$$

after some simplification one obtains:

$$\dot{p_i} = \frac{q}{c}\frac{dx_j}{dt}\frac{\partial A_j}{\partial x_i} - q\frac{\partial\phi}{\partial x_i} \tag{1.91}$$

We will now consider (without the constant c in the denominator) the first piece of Eqn. 1.88 $\vec{v} \times \vec{B}$. Expressed in terms of the vector potential $\vec{A}$ this is simply:

$$\vec{v} \times (\vec{\nabla} \times \vec{A})$$

Looking at this component-wise and using the notation in section 1.9 this is:

$$\epsilon_{lmi} v_m \epsilon_{ijk} \partial_j A_k = \epsilon_{lmi} \epsilon_{ijk} \frac{dx_m}{dt} \frac{\partial A_k}{\partial x_j}$$

where $v_m = \frac{dx_m}{dt}$ has been used. Next we write $\epsilon_{lmi}\epsilon_{ijk} = \epsilon_{ilm}\epsilon_{ijk} = \delta_{lj}\delta_{mk} - \delta_{mj}\delta_{kl}$ and our expression becomes $(\delta_{lj}\delta_{mk} - \delta_{mj}\delta_{kl}) \frac{dx_m}{dt} \frac{\partial A_k}{\partial x_j}$.

Using the cases $l = j$, $m = k$, $k \neq l$, $m \neq j$ and $l \neq j$, $m \neq k$, $k = l$, $m = j$ we can summarize:

$$(\vec{v} \times (\vec{\nabla} \times \vec{A}))_l = (\vec{v} \times \vec{B})_l = \frac{dx_m}{dt} \frac{\partial A_m}{\partial x_l} - \frac{dx_j}{dt} \frac{\partial A_k}{\partial x_j}$$

We can write down the last term on the right hand side of the previous equation as:

$$\frac{dx_j}{dt} \frac{\partial A_l}{\partial x_j} = \frac{dx}{dt} \frac{\partial A_l}{\partial x} + \frac{dy}{dt} \frac{\partial A_l}{\partial y} + \frac{dz}{dt} \frac{\partial A_l}{\partial z}$$

Now one can write the proper derivative of each component of $\vec{A}$ with respect to time as:

$$\frac{dA_l}{dt} = \frac{\partial A_l}{\partial t} + \frac{dx}{dt} \frac{\partial A_l}{\partial x} + \frac{dy}{dt} \frac{\partial A_l}{\partial y} + \frac{dz}{dt} \frac{\partial A_l}{\partial z}$$

$$\therefore \frac{dx_j}{dt} \frac{dA_l}{dx_j} = \frac{dA_l}{dt} - \frac{\partial A_l}{\partial t}$$

$$\therefore \left(\vec{v} \times \vec{B}\right)_l = \frac{dx_k}{dt} \frac{\partial A_k}{\partial x_l} - \left(\frac{dA_l}{dt} - \frac{\partial A_l}{\partial t}\right)$$

In terms of an index i, therefore (repeated indices are summed over this means can choose any symbol to represent the summed index)

$$\therefore \frac{dx_j}{dt} \frac{\partial A_j}{\partial x_i} = \left(\vec{v} \times \vec{B}\right)_i + \frac{dA_i}{dt} - \frac{\partial A_i}{\partial t}$$

Therefore one can substitute for $\frac{dx_j}{dt} \frac{\partial A_j}{\partial x_i}$ in Eqn. 1.91 to obtain:

$$\dot{p}_i = \frac{q}{c} \left\{ (\vec{v} \times \vec{B})_i + \frac{dA_i}{dt} - \frac{\partial A_i}{\partial t} \right\} - q \frac{\partial \phi}{\partial x_i}$$

Since $E = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}$ the last two terms are simply qE_i which means we have obtained:

$$\dot{p}_i = \frac{q}{c}(\vec{v} \times \vec{B})_i + \frac{q}{c}\frac{dA_i}{dt} + qE_i$$

We can isolate the equation of the force on the right hand side by taking $\dot{A}_i = \frac{dA_i}{dt}$ to the left :

$$\dot{p}_i - \frac{q}{c}\dot{A}_i = \frac{q}{c}(\vec{v} \times \vec{B})_i + qE_i$$

So the mechanical momentum (recall $\vec{p}$ and it's components here refer to the canonical momentum) in the presence of an external electromagnetic field has to be modified from $\vec{p} \to \vec{p} - \frac{q\vec{A}}{c}$. For a quantum mechanical equation one can modify the operator $\hat{p}_i$ similarly.

Now recall that $(\frac{E}{c}, \vec{p})$ is a Lorentz 4-vector from which the invariant $\frac{E^2}{c^2} - \vec{p}\cdot\vec{p}$ can be constructed. With the discussed transformation of $\vec{p}$ the quantity $\frac{E^2}{c^2} - \vec{p}\cdot\vec{p}$ changes to $\frac{E^2}{c^2} - \vec{p}\cdot\vec{p} + 2\frac{q}{c}\vec{p}\cdot\vec{A}$ which is clearly no longer Lorentz invariant. It is easy to see that by transforming $\frac{E}{c}$ to $\frac{E}{c} - \frac{q}{c}\phi$ the above quantity will transform to $\frac{E^2}{c^2} - \vec{p}\cdot\vec{p} + 2\frac{q}{c}\vec{p}\cdot\vec{A} - 2\frac{E}{c}\frac{q}{c}\phi$. This makes perfect sense since we have added the potential energy divided by the speed of light to $\frac{E}{c}$. Since $2\frac{q}{c}\vec{p}\cdot\vec{A} - 2\frac{E}{c}\frac{q}{c}\phi$ is simply proportional to the Lorentz scalar product of the 4-vectors $(\frac{E}{c}, \vec{p})$ and $(\phi, \vec{A})$, it is clear that the transformation required can be expressed in terms of the momentum 4-vector and the 4-vector potential is $p^\mu \to p^\mu - \frac{q}{c}A^\mu$ and the same holds true for quantum mechanical operator analogues. Thus the operator $\hat{p_\mu}$ changes to $\hat{p^\mu} - \frac{q}{c}A^\mu$ with $\hat{p^\mu} = (i\hbar\frac{\partial}{\partial x^0}, -i\hbar\vec{\nabla})$.

1.14 Charged particle in an electromagnetic field co-variant expression

Recall the equation for the force on a charged particle in an electromagnetic field.

$$\vec{F} = q\vec{E} + q\frac{\vec{v}}{c} \times \vec{B} \tag{1.92}$$

We will now write down a relativistically covariant expression for the motion of a charged particle magnetic field. We begin by first writing the i^{th} component of equation 1.94

$$F_i = m\frac{dx_i}{dt^2} = qE_i + \epsilon_{ijk}v_jB_k \tag{1.93}$$

where the component indices run from 1 to 3. Recalling the components of the elctromagnetic tensor $F^{\alpha\beta}$ in equation 1.80 we can easily write by examining the tensor's components:

$$F_i = m\frac{dx_i}{dt^2} = qE_i + \epsilon_{ijk} q\frac{v_j}{c} B_k \tag{1.94}$$

where the reader is cautioned about the choice of the same symbol for the tensor and the force-the author hopes that the number of indices will make this unambiguos ! With Roman letters denoting the spatial components of the tensor we have

$$F_i = m\frac{dx_i}{dt^2} = qF^{i0} - q\frac{v_j}{c} F^{ij} \tag{1.95}$$

where only the spatial indices have been used thus far. Recall now the definition of the four velocity and four momenta of a particle :

$$v^\alpha = (\frac{dx^0}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}) = (\gamma c, \gamma\vec{v})$$

and

$$p^\alpha = mv^\alpha = (m\gamma c, \gamma m\vec{v})$$

where of course $p^0 = \frac{E}{c}$ with E here representing the total relativistic energy of the particle. The rate of change of the energy with respect to time is simply the work done on the particle by the electric field, and thus we can write:

$$\frac{cdp^0}{dt} = q\vec{E}\frac{d\vec{x}}{dt} = qE_i v_i = qF^{i0} v_i \tag{1.96}$$

multiplying both sides by $\frac{dx^0}{d\tau}$, recalling also that $dx^0 = cdt$ we obtain:

$$\frac{dp^0}{d\tau} = \frac{q}{c}\vec{E}\frac{d\vec{x}}{d\tau} = qF^{i0}\frac{dv_i}{d\tau} \tag{1.97}$$

using $\frac{dx_j}{dt}$ instead of v_j in equation 1.95 and multiplying by $\frac{dx^0}{d\tau}$ as well and obtain:

$$\frac{dp_i}{d\tau} = qF^{i0}\frac{dx^0}{d\tau} - q\frac{dx_j}{d\tau} F^{ij} \tag{1.98}$$

combining equations 1.97 and and identifying the components of the four momentum we write down the equation of the four-force felt by a charged particle in an electromagnetic field, recall that since $d\tau$ is invariant this is the equation for a four-vector:

$$\frac{dp^\alpha}{d\tau} = \frac{q}{mc} F^{\alpha\beta} \frac{dx_\beta}{d\tau} \tag{1.99}$$

$$\frac{d^2x^\alpha}{d\tau^2} = \frac{q}{mc} F^{\alpha\beta} \frac{dx_\beta}{d\tau} \tag{1.100}$$

In chapters 3 and 4 we will interpret solutions of the Dirac equation travelling backward in time, when viewed by an observer with a normal sense of time, being equivalent to particles of opposite charge. This is seen perhaps most clearly in equation 1.100 where replacing τ with $-\tau$ is equivalent to reversing the charge of the particle [15].

1.15 Natural system of units

The CGS-Gaussian system of units have been used in this text so far, these will however be dropped in chapter 4. The CGS-Gaussian units treat length (in centimeters) mass (in grams) and time (in seconds) as fundamental units, in addition the magnetic field (this changes CGS to CGS-Gaussian) is expressed in the same units as the electric field (for a convenient description of radiation) as we have seen in the Lorentz force law expressed earlier in this chapter. Any other physical quantity is expressed in terms of units of mass, length and time (M, L, T). As an MKS units have units of coulombs for charge, in CGS units these are not used and charge has units $M^{\frac{1}{2}} L^{\frac{3}{2}} T^{-1}$, the electrostatic units (esu) are defined i n terms of these as convenience however charge does not acquire its own fundamental unit as in the MKS system.

Why does the CGS-Gaussian system treat mass, length and time this way ? Because it makes perfect sense for those who deal with the relevant phenomena to do this. Scientists dealing in distances time and weight will find these units very convenient. As physicists about to embark on a study of relativistic quantum mechanics we have to think about what we want to measure and what is the most convenient way to do this. First of all in the relativistic regime we will be dealing with velocities that are substantial fractions of the velocity of light. In quantum mechanics multiples of $\hbar$ become important-this has the unit of action or of angular momentum. Finally we'll be using electrical fields to accelerate fundamental particles and also measuring their energies so that a convenient measure of energy becomes the energy gained by an electron accelerated through a volt-the electron volt. The speed of light, Plancks constant, and the electron volt become important to us since we won't be normally measuring distances or time. As opposed to CGS units which are expressed in term of length mass and time, for us velocity, action and energy become fundamental units. All *other units are expressed in terms of these* in addition we drop factors of $\hbar$ and c everywhere effectively setting them equal to 1. With this convention for units the fine structure constant $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$, with $\hbar = c = 1$ alpha simply is denoted by e (chapter 4).

Chapter 2

Towards a relativistic quantum mechanical wave equation: Dirac Equation

TEST In this chapter a quantum mechanical wave equation compatible with special relativity will be derived. WE will begin by using quatum mechanical energy and momentum operators to replace E and $\vec{p}$ in the Lorentz invariant $E^2 = c^2\vec{p}\cdot\vec{p} + m^2c^4$ to form an operator equation acting on a wave function. This will lead us to the Klein-Gordon equation. However negative energy solutions and a non-positive definite probability density in the continuity equation will be shown to result from non-linearity of the second order time derivative in the operator $\hat{E}^2$. Historically, the negative energy solutions and the non-positive definite probability density were first viewed as serious shortcomings. In time these were re-interpreted in the context of oppositely charged scalar particles, however the percieved shortcomings provided the motivation for the derivation of a quantum mechanical wave-equation consistent with special relativity but with a first order time derivative. Such an equation was derived by Dirac and is known as the Dirac equation. A positive definite probability density is straightforwardly derived from the Dirac equation hence remedying one of the shortcomings of the Klein-Gordon equation. The negative energy ($E < 0$) solutions however remained and were assumed to fill a "sea" of $E < 0$ states. Thus, the state of the vaccum was postulated to be an infinite number of egatively charged particles filling these states. An excitation from any of these states to an $E > 0$ state would leave an unoccupied negative energy "hole" in the sea and appear to an observer as a positively charged particle otherwise identical to an electron, thus a positively charged electron. This conjecture was shown experimentally to be entirely true by Anderson in 1932 with the discovery of the positron and was one of the greatest successes of the Dirac equation. The Dirac equation (as will be shown in this chapter) also *intrinsically* contains spin and accounts for the fine structure and all other relativistic corrections (excluding the Lamb shift) to the hydrogenic spectrum, recall that in the Schrödinger treatment of the hydrogen atom, spin and the relativistic corrections have to be included in by hand. These facts give us confidence that it is the right equation to describe spin $\frac{1}{2}$ particles.

2.1 A first attempt to write down a relativistic quantum mechanical wave equation: The Klein Gordon Equation

An attempt to write down a quantum mechanical wave equation can be made by introducing operators for energy and momenta in the following equation summarizing the relationship between the total relativistic energy and momentum:

$$E^2 = m^2c^4 + c^2\vec{p}\cdot\vec{p} \tag{2.1}$$

We now use the quantum mechanical operator analogues $\hat{E} = i\hbar\frac{\partial}{\partial t}$ and $\hat{p} = -i\hbar\vec{\nabla}$ in place of E and $\vec{p}$ and operate on a wave function $\psi(x,t)$ note that either

$$\hat{E}\psi(x,t) = (m^2c^4 + c^2\hat{p}^2)^{1/2}\psi(x,t) \tag{2.2}$$

or

$$\hat{E}^2\psi(x,t) = (m^2c^4 + c^2\hat{p}^2)\psi(x,t) \tag{2.3}$$

could be an appropriate choice of a quantum mechanical wave equation consistent with special relativity. Let's shoot down one of the choices immediately: the square root in equation 2.2 would mean that we would have to make an infinite expansion in the operator $\hat{p}$. To avoid this so we adopt equation 2.3 as our first candidate for a prototype equation. Recall that $p^\mu = (\frac{E}{c},\vec{p})$ So since $\hat{E} = i\hbar\frac{\partial}{\partial t}$, the operator representing $\frac{E}{c}$ is $\hat{E} = i\hbar\frac{\partial}{\partial(ct)} = i\hbar\frac{\partial}{\partial x^0}$ and so the operator analogue of $\hat{p}^\mu$ can simply be written as:

$$\hat{p}^\mu = i\hbar\partial^\mu = i\hbar(\frac{\partial}{\partial(ct)}, -\vec{\nabla})$$

Thus the prototype equation

$$\hat{E}^2\psi(x,t) = (m^2c^4 + c^2\hat{p}^2)\psi(x,t)$$

becomes

$$\frac{1}{c^2}\frac{\partial^2\psi(x,t)}{\partial t^2} - \nabla^2\psi + \frac{m^2c^2}{\hbar^2}\psi = 0$$

after division by $\hbar^2 c^2$. Using the definition of the d'Alembertian operator $\Box$ we rearrange to obtain the Klein Gordon equation:

$$\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(x,t) \tag{2.4}$$

which has solutions proportional to:

$$e^{-\frac{iEt}{\hbar}+\frac{i\vec{p}\cdot\vec{x}}{\hbar}}\text{with } E = \pm(p^2c^2 + m^2c^4)^{1/2}$$

The negative energy solution above resulting from equation 2.4 was at first considered a problem of the Klein-Gordon equation, the second feature of this equation that was also viewed as a problem was a non-positive definite probability density which we will derive later in this chapter. It should be clear that these issues both have their origins in the second order derivative with respect to time in the Klein Gordon equation–this led to the motivation for the development of the Dirac equation–an equation first order in space and time derivatives.

2.2 On the path to a relativistic quantum mechanical equation: a problem with non-positive definite probability density

To illustrate the second "problem" with the Klein-Gordon equation we will first derive the continuity equation for the conservation of probability from the Schrödinger equation and then do the same for the Klein-Gordon equation. Assuming a *real* potential V, we begin with the Scrödinger equation and its complex conjugate :

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \\ -i\hbar\frac{\partial\psi^*}{\partial t} &= -\frac{\hbar'^2}{2m}\nabla^2\psi^* + V\psi^* \end{aligned} \tag{2.5}$$

multipling the first equation by on the left by ψ^* and the second by ψ and subtracting we obtain after cancelling one power of $\hbar$:

$$\psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t} = \frac{-i\hbar}{2m} \vec{\nabla} \cdot (\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) \tag{2.6}$$

this can easily be rewritten as :

$$\frac{\partial}{\partial t}(\psi^* \psi) + \frac{i\hbar}{2m} \vec{\nabla} \cdot (\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) = 0 \tag{2.7}$$

By identifying $\rho = \psi^* \psi$ as the probability density we must identify

$$\vec{j} = \frac{\hbar}{2mi}(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*)$$

as a probability flux-it should be verified by the reader that it has units of probability per area per time. Then using $\vec{j}$ and ρ above in equation 2.7

$$\frac{\partial}{\partial t}(\rho) + \vec{\nabla} \cdot \vec{j} = 0 \tag{2.8}$$

It is important to note that since ψ is a complex number $\psi^* \psi$ is necessarily positive definite. The same steps can be applied to the Klein-Gordon equation:

$$\begin{aligned} \psi^* \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \psi^* \nabla^2 \psi + \frac{m^2 c^2}{\hbar^2} \psi^* \psi + V \psi^* \psi = 0 \\ \psi \frac{1}{c^2} \frac{\partial^2 \psi^*}{\partial t^2} - \psi \nabla^2 \psi^* + \frac{m^2 c^2}{\hbar^2} \psi \psi^* + V \psi \psi^* = 0 \end{aligned} \tag{2.9}$$

and then by subtracting and rearranging we obtain:

$$\frac{\partial}{\partial t}(\frac{1}{c^2} \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}) + \vec{\nabla} \cdot (\psi \vec{\nabla} \psi^* - \psi * \vec{\nabla} \psi) = 0 \tag{2.10}$$

Multiplying the above equation by $\frac{i\hbar}{2m}$ and using $i = \frac{-1}{i}$ we obtain:

$$\frac{\partial}{\partial t}(\frac{i\hbar}{2mc^2} \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}) + \frac{\hbar}{2mi} \vec{\nabla} \cdot (\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) = 0 \tag{2.11}$$

We identify the probability current $\vec{j} = \frac{\hbar}{2mi}(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*)$ in the light of equations 2.5–2.8, so the probability density is :

$$\rho = \frac{i\hbar}{2mc^2}(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}) \tag{2.12}$$

Clearly $\rho = \frac{i\hbar}{2mc^2}(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t})$, it is no longer guaranteed to be positive definite. It should be obvious to the reader that this behavior of the probability density derived from the Klein-Gordon equation is due to the second-order time derivative in the D'Alembertian operator $\Box$.

As stated in the preamble to this chapter, the non-positive definiteness of ρ was viewed as a problem at first and the Klein-Gordon equation was not considered a good choice for a quantum mechanical wave equation consistent with special relativity. This provided the motivation for the search for an equation linear in the time derivative with a positive definite ρ. Later, with the discovery of charged spin 0 particles (such as charged pions) the negative ρ could be reinterpreted as a charge density. Before the reinterpretation this perceived shortcoming provided the motivation for the search for an equation linear in the time derivative-this led to the Dirac equation. When we cover the Dirac equation and its solutions it will also become obvious that it is succesful in describing spin and relativistic corrections to the hydrogen atom (although we will not derive the hydrogen spectrum). These attributes and the interpretation of the negative energy solutions as spin $\frac{1}{2}$ positrons made it an obvious choice in describing the interactions of electrons and positrons-as we shall see. For now however we follow history, using the "shortcomings" of the Klein-Gordon equation to motivate the search for an equation linear in time.

2.3 Searching for relativistic quantum mechanical equation linear in time

We will now follow the approach of Dirac and find a quantum mechanical wave equation that is linear in time and consistent with special relativity. Let's first summarize what we are looking for:

1. An equation that is linear in the time derivative with a positive definite probability density in the derived continuity equation.
2. An equation that can be used to reproduce the equation

$$\hat{E}^2\psi = c^2\hat{p}\cdot\hat{p}\psi + m^2c^4\psi$$

   with $\hat{E} = i\hbar\frac{\partial}{\partial t}$ and $\hat{p} = -i\hbar\vec{\nabla}$ corresponding to the energy-momentum relation:

$$E^2 = c^2\vec{p}\cdot\vec{p} + m^2c^4$$

3. An equation that is relativistically covariant , i.e. whose *form* is preserved under a Lorentz transformation. Here the word covariant

> is used in a different way-it doesn't refer to the transformation property of a vector-rather it refers to the fact that the form of an equation remains the same from frame to frame. This topic will be dealt with in chapter 3.

We will simply follow the historical approach and search for such an equation. Linearizing the equation $\hat{E}^2\psi = c^2\hat{p}\cdot\hat{p}\psi + m^2c^4\psi$ is equivalent to expressing energy operator as a linear combination of the momenta and rest energy

$$\hat{E}\psi = c\vec{\alpha}\cdot\hat{p}\psi + \beta mc^2\psi$$

where the coefficients $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and β must be determined. We now explicitly introduce the energy and momentum operators and write :

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar c\left(\alpha_1\frac{\partial\psi}{\partial x^1} + \alpha_2\frac{\partial\psi}{\partial x^2} + \alpha_3\frac{\partial\psi}{\partial x^3}\right) + mc^2\beta\psi \tag{2.13}$$

We want to find αs and a β such that when we square this expression we will get simply $\hat{E}^2\psi$ on the left and $\hat{p}\cdot\hat{p}c^2\psi + m^2c^4\psi$ on the right.

In other words the operation :

$$\begin{aligned}-\hbar^2\frac{\partial}{\partial t}(\frac{\partial\psi}{\partial t}) = -\hbar^2c^2&\left(\alpha_1\frac{\partial}{\partial x^1} + \alpha_2\frac{\partial}{\partial x^2} + \alpha_3\frac{\partial}{\partial x^3} - \frac{mc}{i\hbar}\beta\right)\times \\ &\left(\alpha_1\frac{\partial}{\partial x^1} + \alpha_2\frac{\partial}{\partial x^2} + \alpha_3\frac{\partial}{\partial x^3} - \frac{mc}{i\hbar}\beta\right)\psi\end{aligned} \tag{2.14}$$

should yield the operator equation (acting on ψ) analogous to a well known relationship between total relativistic energy and momentum:

$$\hat{E}^2\psi = c^2\hat{p}^2\psi + m^2c^4\psi \tag{2.15}$$

If we carry out the operation shown in equation 2.14 we will obviously obtain cross-terms which do not give us the required relationship between relativistic energy and momentum and so it should be clear that the αs and β are not numbers but operators (in order for the cross terms to dissappear we have to assume non-commutativity of their products). We assume that α and β are matrices, so ψ is a column vector, we denote the components of ψ by ψ_σ each of these must satisfy:

$$-\hbar^2\frac{\partial^2\psi_\sigma}{\partial t^2} = -\hbar^2c^2\nabla^2\psi_\sigma + m^2c^4\psi_\sigma \tag{2.16}$$

This is the expression we would like to obtain by expanding the right hand side of equation 2.14. This will impose requirements on the properties of α and β.

Before we proceed, a clarification: it should be clear that since α and β are matrices, these operate on the components of ψ only. Thus any term such as $\alpha_i\psi$ is in terms of indices $\alpha_{i,\sigma\delta}\psi_\delta$. Since ψ is a column vector this should be easy to understand. Note that the derivative operators "commute" with the α and β matrices and mutiply every element of each matrix them and then operate on each element of ψ. So we proceed and use symbols i and j (where $i \neq j$) to label the spatial indices and obtain:

$$\begin{aligned} -\hbar^2\frac{\partial^2\psi}{\partial t^2} =& -\hbar^2c^2\alpha_i\alpha_i\frac{\partial^2\psi}{\partial x^{i,2}} - \hbar^2c^2(\alpha_i\alpha_j + \alpha_j\alpha_i)\frac{\partial^2\psi}{\partial x^i\partial x^j} \\ & - \frac{\hbar mc^3}{i}(\alpha_i\beta + \beta\alpha_i)\frac{\partial\psi}{\partial x^i} + \beta^2m^2c^4\psi \end{aligned} \tag{2.17}$$

where the explicit indices of the matrices and ψ have been suppressed, for convenience.

By comparing equations 2.16 and 2.17 we come up with the following conditions on the α and β:

1. $\alpha_i\alpha_j + \alpha_j\alpha_i = 0$ (i is understood to be $\neq j$)
2. $\alpha_i\beta + \beta\alpha_i = 0$
3. $\alpha_i^2 = \mathbb{I}$
4. $\beta^2 = \mathbb{I}$ (the identity matrix)

By inspection we know that we can summarize the above in terms of the anti-commutation relations $\{\alpha_i, \beta\} = 0$, $\{\alpha_i, \alpha_j\} = 2\delta_{ij}\mathbb{I}$ and the requirement that $\beta^2 = \mathbb{I}$. The next question is: what is the dimensionality of these matrices? For this we consider the case where $i \neq j$ and begin with

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 0 \implies \alpha_i\alpha_j = -\alpha_j\alpha_i$$

multiplying on the left by α_j we obtain $\alpha_j\alpha_i\alpha_j = -\alpha_j^2\alpha_i$ using $\alpha_j^2 = \mathbb{I}$ we get $\alpha_j\alpha_i\alpha_j = -\alpha_i$, taking the trace of both sides:

$$Tr(\alpha_j\alpha_i\alpha_j) = -Tr(\alpha_i)$$

by the cyclic property of the trace:

$$Tr(\alpha_j\alpha_i\alpha_j) = Tr(\alpha_j^2\alpha_i) = -Tr(\alpha_i)$$

and so finally we obtain

$$Tr(\alpha_i) = -Tr(\alpha_i) \implies 2Tr(\alpha_i) = 0$$

So the α_i have trace$=0$. Similarly one can show that β is also traceless.

Since the trace of any matrix is also equal to the sum of its eigenvalues, the α_i and β must be *even*-dimensional matrices since only an even number of 1 and -1 can sum to zero. Let us now summarize what we have learnt so far:

1. The matrices α_i, β are traceless
2. They are even dimensional with eigenvalues ± 1
3. Since their eigenvalues are real it follows that they must be Hermitian.
4. Their squares should equal the identity matrix $\mathbb{I}$.
5. They should anti-commute .

At first glance one might assume that the Pauli matrices satisfy these conditions, however there are 3 Pauli matrices and we need a fourth matrix (β) as well. Note that the three Pauli matrices (proof follows) and the 2×2 identity matrix form a basis for all two dimensional Hermitian matrices. In other words any 2×2 Hermitian matrix can be written as a linear combination of the 3 Pauli matrices and $\mathbb{I}$. Proof follows:

Proof that σ_i and $\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ form a complete basis for all 2×2 Hermitian matrices. Let

$$\begin{pmatrix} u & v \\ v^* & w \end{pmatrix} \text{ represent your most general } 2\times 2 \text{ Hermitian matrix}$$

of course u and w are real.

We write this as a linear combination with real coefficients for the Pauli matrices and the identity matrix:

$$\begin{pmatrix} u & v \\ v^* & w \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$u = a + d$, $v = c + ib$, $v^* = c - ib$, $w = d - a$.

We see that a, b c and d can always be solved for with: $d = \frac{u+v}{2}$, $a = \frac{u-w}{2}$, $c = \Re(v)$ and $b = \Im(v)$, $\therefore$ $(\sigma_i, \mathbb{I})$ form a complete basis for all 2×2 Hermitian matrices. QED.

We need a fourth traceless, Hermitian that anti-commutes with three Pauli matrices, In an attempt to get a fourth such matrix we can't include the identity matrix in the linear combination since the resulting matrix will trace$\neq 0$. So we must look only at linear combinations of only the

Pauli matrices to provide. We can't pick a linear combination that includes the three Pauli matrices since it is then impossible to satisfy the anti-commutation relations between 4 matrices, we can't pick two of three for the same reason, and the only choice that is left is a multiple of one Pauli matrix. It should be clear that we cannot satisfy the requirements on α_i and β in two dimensions.

The next even dimension to consider is 4 and the reader should verify that the following choice of 4×4 matrices satisfy all the conditions on the α and β matrices:

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \tag{2.18}$$

where each of the entries σ_k, $\mathbb{I}$m in the above are the 2×2 Pauli and identity matrices respectively. Written explicitly these are:

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{2.19}$$

Now that we have found an appropriate choice of α and β we rewrite our prototype equation explicity inserting the indices for the column "vector" ψ and the α and β matrices . The prototype equation for the σth component ψ (which we now know is a 4-component object) is now:

$$i\hbar \frac{\partial \psi_\sigma}{\partial t} = -i\hbar c \alpha_{k,\sigma\rho} \frac{\partial \psi_\rho}{\partial x^k} + mc^2 \beta_{\sigma\tau} \psi_\tau \tag{2.20}$$

Before deriving the continuity equation for this equation, we will write down the Hermitian conjugate of equation 2.20, note that the operation of Hermitian conjugation on a column vector will simply takes it's complex conjugate element by element and turns it to the corresponding row vector. Thus we have

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \quad \text{and} \quad \psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \tag{2.21}$$

Note that the operation of matrix multiplication is defined as: $\chi_\sigma = \alpha_{k,\sigma\rho}\psi_\rho$ for the column vector and $\chi_\sigma^\dagger = \psi_\rho^\dagger \alpha_{k,\rho\sigma}$ for the Hermitian conjugate. Now $(\alpha_k)^\dagger = \alpha_k$ and $(\beta)^\dagger = \beta$ since they are Hermitian, keeping this in mind we obtain for the Hermitian conjugate of equation 2.20:

$$-i\hbar\frac{\partial\psi_\sigma^*}{\partial t} = i\hbar c\frac{\partial\psi_\rho^*}{\partial x^k}\alpha_{k,\rho\sigma} + mc^2\psi_\tau^*\beta_{\tau\sigma} \tag{2.22}$$

Now take the scalar product of equation 2.22 (on the left) with $\psi^\dagger$ and equation 2.20 on the right by ψ, using explicit indices for now we have:

$$i\hbar\psi_\sigma^*\frac{\partial\psi_\sigma}{\partial t} = -i\hbar c\psi_\sigma^*\alpha_{k,\sigma\rho}\frac{\partial\psi_\rho}{\partial x^k} + mc^2\psi_\sigma^*\beta_{\sigma\tau}\psi_\tau \tag{2.23}$$

$$-i\hbar\frac{\partial\psi_\sigma^*}{\partial t}\cdot\psi_\sigma = i\hbar c\frac{\partial\psi_\sigma^*}{\partial x^k}\alpha_{k,\sigma\rho}\psi_\rho + mc^2\psi_\sigma^*\beta_{\sigma\tau}\psi_\tau \tag{2.24}$$

Subtracting equation 2.23 from equation 2.24, suppressing the explicit indices of summation and using $\psi^\dagger$ and ψ and re-arranging we obtain:

$$-i\hbar\frac{\partial\psi^\dagger}{\partial t}\cdot\psi - i\hbar\psi^\dagger\frac{\partial\psi}{\partial t} = i\hbar c\frac{\partial\psi^\dagger}{\partial x^k}\alpha_k\psi + i\hbar c\psi^\dagger\alpha_k\frac{\partial\psi}{\partial x^k} \tag{2.25}$$

this can easily be rewritten as:

$$\frac{\partial}{\partial t}(\psi^\dagger\psi) + \frac{\partial}{\partial x^k}(c\psi^\dagger\alpha_k\psi) = 0 \tag{2.26}$$

The expression in equation 2.26 is easily identified as a continuity equation with $\psi^\dagger\psi$ as the probability density ρ and $c\psi^\dagger\alpha_k\psi$ as $\vec{\jmath}$ of the probability current and note that the k^{th} component of $\vec{\jmath}$ is associated with α_k. Its obvious that ρ is a probability density and is positive definite. It is easily seen that the units of $\vec{j}$ are probability per area per time-as expected. Note that equation 2.26 can be rewritten :

$$\frac{\partial}{\partial(t)}(\psi^\dagger\psi) + \frac{\partial}{\partial x^k}(c\psi^\dagger\alpha_k\psi) = 0 \tag{2.27}$$

Dividing equation 2.27 by c to obtain the partial derivative with respect to $x_0 = ct$, identifying ∂_μ we can re-write equation 2.26 compactly as:

$$\partial_\mu j^\mu = 0 \tag{2.28}$$

where we have defined:

$$j^\mu = (c\rho, \vec{j}) = (c\psi^\dagger\psi, c\psi^\dagger\alpha_k\psi) \tag{2.29}$$

Some liberty has been taken here using the notation j^μ, however in the next chapter we will show that it is actually a contravariant vector rigorously. Note that the spatial part of j^μ has units of a probability per area per time and can be used to represent a flux of particles. Suffice to say at this point we have identified a positive definite probability density and a current.

2.4 Solutions of the Dirac equation for a particle at rest

To solve the Dirac equation for a particle at rest we begin with the prototype Dirac equation 2.23 with the spatial components set to zero

$$i\hbar\frac{\partial\psi}{\partial t} = i\hbar\beta\frac{\partial\psi}{\partial(t)} = mc^2\psi$$

writing this out explicitly:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = mc^2\begin{pmatrix}\psi_1\\ \psi_2\\ -\psi_3\\ -\psi_4\end{pmatrix}$$

These are 4 trivial differential equations and are easily solved:

$$\begin{aligned}
i\hbar\frac{\partial\psi_1}{\partial t} &= mc^2\psi_1 \quad \rightarrow \psi_1 = e^{-i\frac{mc^2t}{\hbar}}\\
i\hbar\frac{\partial\psi_2}{\partial t} &= mc^2\psi_2 \quad \rightarrow \psi_2 = e^{-i\frac{mc^2t}{\hbar}}\\
-i\hbar\frac{\partial\psi_3}{\partial t} &= mc^2\psi_3 \quad \rightarrow \psi_3 = e^{i\frac{mc^2t}{\hbar}}\\
-i\hbar\frac{\partial\psi_4}{\partial t} &= mc^2\psi_4 \quad \rightarrow \psi_4 = e^{i\frac{mc^2t}{\hbar}}
\end{aligned} \tag{2.30}$$

The four solutions written explicitly are:

$$\psi_1(x) = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} e^{-i\frac{mc^2t}{\hbar}}, \psi_2(x) = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} e^{-i\frac{mc^2t}{\hbar}} \tag{2.31}$$

these are the positive energy solutions with $E = +mc^2$. The third and fourth are negative energy solutions with $E = -mc^2$:

$$\psi_3(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{+i\frac{mc^2 t}{\hbar}}, \psi_4(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{+i\frac{mc^2 t}{\hbar}} \tag{2.32}$$

So although we are rid of the negative probability density we are still stuck with a negative energy solutions. We also have two negative solutions for every energy eigenvalue, it is tempting to interpret these as spin. We now construct the 4×4 analogue of the Pauli matrix σ_z which we denote by Σ_z

$$\Sigma_z = \begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} \tag{2.33}$$

written explicitly this is

$$\Sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.34}$$

it is trivial to see that the stationary solutions in equation 2.32 are eigenvectors of $S_z = \frac{\hbar}{2}\Sigma_z$ with eigenvalues $\pm\frac{\hbar}{2}$. This is a strong hint that these extra degrees of freedom are related to spin. We shall confirm this in this chapter and the next. For the remainder of this section we will assume that the Dirac equation describes electrons, with this assumption made we now proceed to interpret the $E < 0$ solutions.

Although we have solved the issue of a non-positive definite probability density we do have to interpret the physical meaning of the negative energy $(E < 0)$ solutions. If $E < 0$ states are available to electrons with $E > 0$ then in time (and in a very short time) these would all radiate a photon and fall into these states. To remedy this paradox Dirac postulated that all $E < 0$ states are already filled with electrons and of course the Pauli exclusion principle prevents more than 1 from filling a given state. Thus a sea of filled $E < 0$ states with negative charge *establishes our reference for all other observations*. The largest energy corresponding to an electron filling a $E < 0$ state is then simply the rest energy$\times -1 = -mc^2$. As in atoms where electrons transit from one state to another whilst absorbing or emitting photons, Dirac postulated that electromagnetic radiation (this would have to be more than 1 photon as we'll see in a later chapter) could impart enough energy to an electron occupying a $E < 0$ state to allow it to transit to a $E > 0$ state, leaving an unoccupied state in the negative energy sea. Let us now go through the consequences of this: for the purposes of illustration let us assume the negative energy sea contains a large number N of $E < 0$ electron states and thus charge $-Ne$. Removing one negatively charged particle from this means we now have $-Ne - (-e) = -Ne + e$ the charge of the sea (which is our reference and we dont observe it) plus another $+e$. If the total energy of the the negative energy sea is $\sum_i^N E_i$ and we remove a $E < 0$ electron from it we obtain $\sum_i^N E_i - (-E) = \sum_i^N E_i + E$ once again the first term is our reference. It should be clear that an unoccupied $E < 0$ state appears to us as an oppositely charged particle of positive energy and opposite charge. A similar argument can be carried out for spin and for momentum. Thus an unoccupied negative energy state of charge $-e$, energy $-E$, momentum p and spin s appears to us as a particle of charge e, energy E momentum $-p$ and spin $-s$. It is also possible that if one of the negative energy states were empty then an electron with positive energy could occupy this giving up the energy difference as a radiated photon-in this case a hole and and electron would dissappear with the appearance of electromagnetic radiation. Given that the highest negative energy state is $-mc^2$ and the lowest positive energy state is mc^2 any of these processes would require the absorption or emission of a photon of energy at least equal to $E = 2mc^2$.

The processes we have just described in this simple picture are the familiar pair creation and pair annhilation of electrons and positrons. Dirac's conjecture was striking since in 1932 a positively charged spin $\frac{1}{2}$ particle of equal mass to the electron (the positron) was discovered. Mathematically positrons are equivalent to electrons of negative energy travelling backward in space and time-this equivalence will be demonstrated in the chapter 3 and will be used to describe positrons.

This picture is quite unsatisfactory and unappealing, also contains inconsistencies for example all bosons would all fall into one negative energy state ! The correct treatment of positrons and electrons lies in quantum field theory where electrons and positrons can be created from and annihilated into the vaccum. The negative energy sea of Dirac was the first description of the quantum mechanical vacumm.

2.5 The success of the Dirac Equation in incorporating spin and relativistic corrections

In this section we will demonstrate that the Dirac equation contains a description of spin $\frac{1}{2}$ particles and in the presence of central potential and an external field all the relativistic corrections to the hydrogen spectrum. This was mentioned at the end of section 2.2 of this chapter. The reader is reminded that for the Schrödinger equation, spin and the relativistic corrections all have to be introduced by hand [11]. We will first demonstrate this for the Schrödinger equation and then also demonstrate that the Schrödinger equation conserves *orbital* angular momentum in a central potential. We will then demonstrate that a description of spin $\frac{1}{2}$ particles is intrinsically contained in the Dirac equation and that the Dirac equation conserves *total* angular momentum in a central potential. We will also demonstrate that the relativistic corrections to the Hydrogen atom are contained in the Dirac equation however a solution of the Dirac equation for the hydrogen atom will however *not* be provided, the reader is referred to other texts for this [8] [7] [10]. We will begin by demonstrating how spin is introduced arbitrarily into the Schrödinger equation. Let us now pretend that all we know is the Schrödinger equation and think that $\psi(x)$ is a one component wave function. We are then confronted with the results of the Stern-Gerlach experiment and so must now account for the presence of a magnetic dipole produced by an intrinsic angular momentum of the electron that takes on only two discrete values. A set of matrices, that are 2×2 with eigenvalues $\pm\frac{\hbar}{2}$ following the algebra of angular momentum seem to be the natural choice. We simply remind the reader that

$$[s_i, s_j] = i\hbar\epsilon_{ijk}s_k \tag{2.35}$$

with the definition

$$s_i = \frac{\hbar}{2}\sigma_i \tag{2.36}$$

where the σ_i satisfy the following commutation and anti-commutation relations:

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.37}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \tag{2.38}$$

form an appropriate choice-the σ_i are of course the well known Pauli matrices, these are assumed to operate on a two-component wave function. We now assume that the Schrödinger equation (without a potential for simplicities sake):

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x) = i\hbar\frac{\partial\psi}{\partial t} \tag{2.39}$$

contains this two component wave function, thus $\psi(x) = \psi_\alpha(x)$ with $\alpha = 1, 2$. We however will suppress the component index for convenience. We now ask: what modification should be made to the Schrödinger equation so that it reduces to equation. (2.39) in the absence of a magnetic field and manifests an interaction of

$$-\vec{\mu}\cdot\vec{B} \tag{2.40}$$

when a magnetic field is switched on ? We begin by rewriting the Schrödinger equation as

$$\frac{\hat{p}^2}{2m}\psi(x) = i\hbar\frac{\partial\psi(x)}{\partial t} \tag{2.41}$$

with $\hat{p} = -i\hbar\vec{\nabla}$ and *arbitrarily* make the change $\hat{p} \to \vec{\sigma}\cdot\hat{p}$ in equation (2.41), noting that ψ is now a two component object. Using the anti-commutation relations (equation 2.38) we obtain:

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{2.42}$$

Lets look at (2.41) in the light of the change $\hat{p} \to \vec{\sigma}\cdot\hat{p}$, using indices to illustrate the components this is $\sigma_i\hat{p}_i$ and so we write:

$$(\sigma_i\hat{p}_i)^2 = (\sigma_i\hat{p}_i)(\sigma_j\hat{p}_j) = \sigma_i\sigma_j\hat{p}_i\hat{p}_j = (\delta_{ij} + i\epsilon_{ijk}\sigma_k)\hat{p}_i\hat{p}_j \tag{2.43}$$

where equation 2.42 has been used. In vector notation this is simply:

$$\hat{p}^2\mathbb{I} + i(\hat{p}\times\hat{p})\cdot\sigma = \hat{p}^2\mathbb{I} + 0 = \hat{p}^2\mathbb{I} \tag{2.44}$$

thus the arbitrary substitution $\hat{p} \to \vec{\sigma}\cdot\hat{p}$ and the assumption that the wave function has two components leaves the Schrödinger equation unchanged. In the presence of a magnetic field $\vec{B} = \vec{\nabla}\times\vec{A}$ we must make the modification $\hat{p} \rightarrow \hat{p} - \frac{e\vec{A}}{c}$ (see Chapter 1) and with this modification the Schrödinger equation becomes

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{1}{2m}(\vec{\sigma}\cdot\left(\hat{p} - \frac{e\vec{A}}{c}\right))^2\psi(x) \tag{2.45}$$

Looking at each component of the operator $(\vec{\sigma}\cdot\left(\hat{p} - \frac{e\vec{A}}{c}\right))^2$ we can write down:

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= \frac{1}{2m}\sigma_i\sigma_j\left(-i\hbar\partial_i - \frac{e}{c}A_i\right)\left(-i\hbar\partial_j - \frac{e}{c}A_j\right)\psi \\ &= \frac{1}{2m}\left(\delta_{ij} + i\epsilon_{ijk}\sigma_k\right)\left(-i\hbar\partial_i\frac{e}{c} - A_i\right)\left(-i\hbar\partial_j - \frac{e}{c}A_j\right)\psi \end{aligned} \tag{2.46}$$

simplifying equation 2.46 and noting that $i\epsilon_{ijk}\sigma_k\frac{ie\hbar}{2mc}\partial_i A_j\psi = i\sigma_k(\nabla\times\vec{A})_k$ is simply $-\frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B}$ we obtain

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\frac{1}{2m}(\hat{p} - \frac{e\vec{A}}{c})^2 - \frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B}\right)\psi \tag{2.47}$$

the term $-\frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B}$ can be written as the negative of the scalar product of a magnetic moment with the applied B field:

$$-\frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B} = -\frac{e}{mc}\vec{s}\cdot\vec{B} = -\vec{\mu}\cdot\vec{B} \tag{2.48}$$

Note that $\vec{\mu}$ is usually written in units of the Bohr magneton $\mu_B = \frac{e\hbar}{2mc}$ in this case therefore we can write

$$\vec{\mu} = g\mu_B\hat{s} \tag{2.49}$$

where $\hat{s}$ is a unit vector in direction of the spin and g is the g-factor and is equal to 2. Note that including a potential in the equation (2.47) gives us the standard form for the Pauli two-component equation-a scalar potential $V = e\Phi(|\ \vec{x}\ |)$ has also been introduced:

$$\left(\frac{1}{2m}\left(\hat{p} - \frac{e\vec{A}}{c}\right)^2 - \frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B} + V\right)\psi = i\hbar\frac{\partial\psi}{\partial t} \tag{2.50}$$

We have demonstrated to the reader that spin must be introduced *by hand* into the Schrödinger equation by changing $\hat{p} \to \vec{\sigma}\cdot\hat{p}$-a completely arbitrary construct. We will now write down the Schrödinger equation without spin and introduce a central potential $V(r) = V(|\vec{x}|)$ and demonstrate that it conserves *orbital* angular momentum. This is a preamble to demonstrating that the Dirac equation will conserve *total* angular momentum and hence that it contains intrinsically the spin angular momentum of a spin $\frac{1}{2}$ particle. The Schrödinger equation with a central potential becomes:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \tag{2.51}$$

where V is understood to be a function of $r = |\vec{x}|$. Recall that the central angular momentum operator is $\vec{x} \times \hat{p}$ writing the k^{th} compoment this is: $-i\hbar\epsilon_{klm}x_l\partial_m$. We will now show that the orbital angular momentum is a conserved quantity i.e. its commutator with the Hamiltonian ($H = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$) is zero. We write this down:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(|\vec{x}|), \vec{x} \times (-i\hbar\vec{\nabla})\right] = 0 \tag{2.52}$$

obviously, it suffices to proove the two relations:

$$[\epsilon_{klm}x_l\partial_m, \partial_i\partial_i] = 0 \tag{2.53}$$

$$[\epsilon_{klm}x_l\partial_m, V(|\vec{x}|)] = 0 \tag{2.54}$$

Dealing with (2.53) first; we operate the commutator on ψ and use the product rule:

$$\epsilon_{klm}x_l\partial_m(\partial_i\partial_i\psi) - \epsilon_{klm}\partial_i\partial_i(x_l\partial_m\psi)$$

expanding out the second term:

$$\begin{aligned}
&-\partial_i\epsilon_{klm}(\partial_i x_l\partial_m\psi + \epsilon_{klm}x_l\partial_i\partial_m\psi) \\
&= -(\epsilon_{klm}\delta_{il}\partial_i\partial_m\psi_i + \epsilon_{klm}\delta_{il}\partial_i\partial_m\psi) - \epsilon_{klm}x_l\partial_i\partial_i\partial_m\psi
\end{aligned}$$

The term within the brackets is simply $2\vec{\nabla}\times\vec{\nabla}\psi$ which is zero, the second term simply cancels $\epsilon_{klm}x_l\partial_m\partial_i\partial_i\psi$ giving zero for Equation 2.53 that is

$[\epsilon_{klm}x_l\partial_m, \partial_i\partial_i] = 0$. Now onto equation 2.54, we operate the commutator on ψ to obtain:

$$[\epsilon_{klm}x_l\partial_m, V(|\vec{x}|)] = \epsilon_{klm}x_l\partial_m(V\psi) - V\epsilon_{klm}x_l\partial_m\psi$$

$$= \epsilon_{klm}x_l\partial_m(V)\psi + \epsilon_{klm}x_l V\partial_m\psi - V\epsilon_{klm}x_l\partial_m\psi$$

the last two terms cancel, and we are left with

$$\epsilon_{klm}x_l\partial_m(V)\psi$$

Note $\partial_m V = \partial_m V(|\vec{x}|) = \frac{x_m}{|\vec{x}|}\frac{dV}{d|\vec{x}|}$ using $|\vec{x}| = (x_j x_j)^{\frac{1}{2}}$
and so we have:

$$\epsilon_{klm}\frac{x_l x_m}{|\vec{x}|}\frac{dV}{d|\vec{x}|}$$

now in the piece $\epsilon_{klm}x_l x_m$ products of each component of $\vec{x}$ occur with odd and even permutations of the indices-in other words $\vec{x}\times\vec{x} = 0$ and hence $[\epsilon_{klm}x_l\partial_m, V(|\vec{x}|)] = 0$. We have shown by the inclusion of a central potential to the Schrödinger equation describes a system in which orbital angular momentum is conserved. We note that this implies that spin orbit interactions that break the degeneracy in the hydrogen atom energy must be introduced by hand into the Schrödinger equation leaving the total angular momentum $\vec{j} = \vec{l} + \vec{s}$ as the good quantum number. We will now show that the Dirac equation conserves the *total* angular momentum $\vec{j} = \vec{l} + \vec{s}$ i.e. both the spin and orbital parts.

To consider the case where a central potential is included we introduce a 4-vector potential whose zeroth component is a central scalar potential i.e. $\Phi(|\vec{x}|)$, so we have $A^\mu = (\Phi(|\vec{x}|), \vec{A})$. The Dirac equation is then modified using the recipe $\hat{p}^\mu \to \hat{p}^\mu - \frac{e}{c}A^\mu$. Since our immediate purpose is to simply consider the case of a central scalar potential we don't consider the spatial part of A^μ ($\vec{A}$ -or the vector potential) and write down

$$i\hbar\frac{\partial}{\partial t}\psi = (-ic\hbar\vec{\alpha}\cdot\vec{\nabla} + V(|\vec{x}|) + \beta mc^2)\psi \tag{2.55}$$

the Dirac equation 2.55 here is written in the *Hamiltonian form* (in the same way as in the Schrödinger equation is here $i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$):

$$i\hbar\frac{\partial}{\partial t}\psi = \hat{H}_{DIRAC}\psi = (-ic\hbar\vec{\alpha}\cdot\vec{\nabla} + V(|\vec{x}|) + \beta mc^2)\psi \tag{2.56}$$

where $-e\Phi(x)$ has now been replaced by $V(\mid \vec{x} \mid)$ for convenience. Similar to our procedure for the Schrödinger equation we must now calculate the commutator $[\hat{H}_{DIRAC}, \vec{x}\times\hat{p}]$ and see if the orbital angular momentum is conserved by a Dirac particle. We have already shown that the commutator of the angular momentum and the potential $[\vec{x}\times\hat{p}, V(|\vec{x}|)] = 0$. The first piece is easily seen to be zero, we consider the second piece:

$$[\hat{H}_{DIRAC}, \vec{x}\times\hat{p}] = [-ic\hbar\vec{\alpha}\cdot\vec{\nabla}, \vec{x}\times\hat{p}]$$

$$= (-ci\hbar)(-i\hbar)[\vec{\alpha}\cdot\vec{\nabla}, \vec{x}\times\vec{\nabla}]$$

The i^{th} component of the commutator then is:

$$[\hat{H}_{DIRAC}, \vec{x}\times\hat{p}]_i = (-ic\hbar)(-i\hbar)[\alpha_l\partial_l, \epsilon_{ijk}x_j\partial_k]$$

Operating $[\alpha_l\partial_l, \epsilon_{ijk}x_j\partial_k]$ on ψ we obtain:

$$\alpha_l\partial_l(\epsilon_{ijk}x_j\partial_k\psi) - \epsilon_{ijk}x_j\partial_k(\alpha_l\partial_l\psi)$$

$$= \alpha_l\epsilon_{ijk}\delta_{lj}\partial_k\psi + \alpha_l\epsilon_{ijk}x_j\partial_k\partial_l - \alpha_l\epsilon_{ijk}x_j\partial_k\partial_l\psi$$

the first and third terms on the right cancel and we are left with:

$$\alpha_l\epsilon_{ijk}\delta_{lj}\partial_k\psi = \epsilon_{ijk}\alpha_j\partial_k\psi \tag{2.57}$$

Thus we have the result:

$$[\hat{H}_{DIRAC}, \vec{x}\times\hat{p}]_i = [\hat{H}_{DIRAC}, \vec{L}]_i = (-ic\hbar)(-i\hbar)\epsilon_{ijk}\alpha_j\partial_k\psi \tag{2.58}$$

This commutator is clearly not zero, so the orbital angular momentum is not conserved. We will now show that it is actually the total angular momentum that is conserved. Recall the commutation relations:

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.59}$$

and the definition of the α matrices:

$$\alpha_j = \left(\begin{array}{c|c} 0 & \sigma_j \\ \hline \sigma_j & 0 \end{array}\right) \tag{2.60}$$

as in section 2.4, we define Σ_i, the 4 dimensional analogues of Pauli matrices which have the same algebra as their 2 dimensional counterparts

$$\Sigma_i = \left(\begin{array}{c|c} \sigma_i & 0 \\ \hline 0 & \sigma_i \end{array}\right) \tag{2.61}$$

and assume that $S_i = \frac{\hbar}{2}\Sigma_i$ represents the spin angular momentum for a Dirac particle. By using equation 2.58 we can write:

$$\begin{aligned}&[\hat{H}_{DIRAC}, \vec{x}\times\hat{p} + \frac{\hbar}{2}\Sigma]_i = \\ &(-ic\hbar)(-i\hbar)\epsilon_{ijk}\alpha_j\partial_k + (-ic\hbar)(\frac{\hbar}{2})[\alpha_m\partial_m, \Sigma_n]\end{aligned} \tag{2.62}$$

It is easy to verify that:

$$[\alpha_m, \Sigma_n] = \left(\begin{array}{c|c} 0 & [\sigma_m, \sigma_n] \\ \hline [\sigma_m, \sigma_n] & 0 \end{array}\right) = 2i\epsilon_{imn}\left(\begin{array}{c|c} 0 & \sigma_i \\ \hline \sigma_i & 0 \end{array}\right) = 2i\epsilon_{imn}\alpha_i \tag{2.63}$$

Using equation 2.63 the right hand side of equation 2.62 is now

$$(-ic\hbar)(-i\hbar)\epsilon_{ijk}\alpha_j\partial_k + (-ic\hbar)(\frac{\hbar}{2})2i\epsilon_{imn}\alpha_i\partial_m$$

permuting some of the indices we obtain finally:

$$\begin{aligned}&[\hat{H}_{DIRAC}, \vec{x}\times\hat{p} + \frac{\hbar}{2}\Sigma]_i \\ &= (-ic\hbar)(-i\hbar)\epsilon_{ijk}\alpha_j\partial_k + (-ic\hbar)(i\hbar)\epsilon_{imn}\alpha_i\partial_m \\ &= (-c\hbar^2)\epsilon_{ijk}\alpha_j\partial_k + (c\hbar^2)\epsilon_{nim}\alpha_i\partial_m = 0\end{aligned} \tag{2.64}$$

So it's easy to see that the *total* angular momentum $\vec{J} = \vec{x}\times\hat{p} + \frac{\hbar}{2}\vec{\Sigma} = \vec{L} + \vec{S}$ is conserved, neither spin nor orbital angular momenta are conserved on their own. We've seen that the Dirac equation contains spin intrinsically unlike the Schrödinger equation where it has to be put in by hand. The hydrogen atom spectrum for the Dirac equation is calculated in several excellent texts [8] [7] [10] [12] but not here. We'll however demonstrate how the Pauli equation follows from the non-relativistic limit of the Dirac equation and also show that the Dirac equations contains all the relativistic corrections to the hydrogen atom (spin-orbit etc). To begin we write down the Dirac equation in the presence of E-M coupling, considering both a vector and scalar potential

$$i\hbar\frac{\partial}{\partial t}\psi = (-ic\hbar\vec{\alpha}\cdot(\vec{\nabla} - e\vec{A}) + V(|\vec{x}|) + \beta mc^2))\psi \tag{2.65}$$

We now split the wave function ψ into a 2-component wave function $\psi = \begin{pmatrix}\phi \\ \chi\end{pmatrix}$ (one will be found to be negligible in the non-relativistic limit). Note the normalization of ψ is given by

$$\int d^3x\psi^\dagger\psi = \int d^3x(\phi^\dagger\phi + \chi^\dagger\chi) = 1 \tag{2.66}$$

so that it is ψ that is normalized to unity and not its two components. Thus now $\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$ substituting this and using $\gamma^0\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}$ explicitly we obtain the equation:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} eV & \vec{\sigma}\cdot(c\hat{p} - e\vec{A}) \\ \vec{\sigma}\cdot(c\hat{p} - e\vec{A}) & eV - 2mc^2 \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.67}$$

this gives us two coupled differential equations for ϕ and χ:

$$\begin{aligned} i\hbar\frac{\partial\phi}{\partial t} &= E\phi = eV\phi + \vec{\sigma}\cdot(c\hat{p} - e\vec{A})\chi \\ i\hbar\frac{\partial\chi}{\partial t} - E\chi &= \vec{\sigma}\cdot(c\hat{p} - e\vec{A})\phi + (eV - 2mc^2)\chi \end{aligned} \tag{2.68}$$

where $\hat{p}$ is the operator $-i\hbar\vec{\nabla}$. If we view equation (2.68) in the non-relativistic limit we can assume that E, $eV << mc^2$ and simply the second equation in (2.68) as:

$$\chi = \frac{\vec{\sigma}\cdot(c\hat{p} - e\vec{A})}{2mc^2}\phi = \frac{\vec{\sigma}\cdot\left(\beta - \frac{e\vec{A}}{c}\right)}{2mc}\phi \tag{2.69}$$

In equation 2.69 the action of the operator $\hat{p}$ has been used to extract the momentum eigenvalue and the velocity $\vec{\beta}$ is simply $\frac{\vec{p}}{m}$. Note that χ is the smaller component of ψ by order $\gamma \mid \vec{\beta} \mid \approx \frac{|\vec{p}|}{mc}$. Substituting for χ in the first expression of equation 2.68 we obtain:

$$\begin{aligned} i\hbar\frac{\partial\phi}{\partial t} &= eV\phi + \frac{\vec{\sigma}\cdot\left(\hat{p} - \frac{e\vec{A}}{c}\right)\vec{\sigma}\cdot\left(c\hat{p} - \frac{e\vec{A}}{c}\right)\phi}{2mc} \\ &= eV\phi + \frac{1}{2m}\vec{\sigma}\cdot\left(\hat{p} - \frac{e\vec{A}}{c}\right)\vec{\sigma}\cdot\left(\hat{p} - \frac{e\vec{A}}{c}\right)\phi \end{aligned} \tag{2.70}$$

Comparing equation (2.70) to 2.45 and then to (2.50) demonstrates clearly that by taking the non-relativistic limit of the Dirac equation in the presence of a magnetic field we have recovered the Pauli equation.

We will now demonstrate that by taking into account first order relativistic corrections to the Dirac equation in the presence of a central potential eV we can generate all the terms that are inserted by hand into the Schrödinger equation to explain the fine structure of hydrogen. The Dirac equation thus contains these intrinsically–just like it contains the fermion spin. We begin by switching off the external magnetic field by setting $\vec{A}=0$. Since the energy now contains a relativistic correction we change E_{NR} to E and rewrite equation 2.68:

$$i\hbar\frac{\partial\phi}{\partial t} = E\phi = eV\phi + \vec{\sigma}\cdot(c\hat{p})\chi \tag{2.71}$$

$$i\hbar\frac{\partial\chi}{\partial t} = E\chi = \vec{\sigma}\cdot(c\hat{p})\phi + (eV - 2mc^2)\chi \tag{2.72}$$

using equation (2.72) we write χ in terms of ϕ, this time E and eV are retained, since although still small compared to $2mc^2$ we do want to use them to calculate relativistic corrections. Doing this we obtain:

$$\chi = \frac{\vec{\sigma}\cdot\hat{p}}{2mc\left(1+\frac{E}{2mc^2}-\frac{eV}{2mc^2}\right)}\cdot\phi \tag{2.73}$$

Using 2.73 in equation 2.71 we obtain:

$$i\hbar\frac{\partial\phi}{\partial t} = E\phi = eV\phi + \frac{(\vec{\sigma}\cdot\hat{p})}{(2m)}\frac{1}{(1+\frac{E}{2mc^2}-\frac{eV}{2mc^2})}(\vec{\sigma}\cdot\hat{p})\phi \tag{2.74}$$

We note however that equation (2.74) describes only *one* component of the Dirac equation, we have dropped χ from consideration without accounting for the fact that we must consider the normalization of $\binom{\phi}{\chi}$ i.e $\int d^3x(\phi^\dagger\phi + \chi^\dagger\chi)$. Using the approximation $\chi = \frac{\vec{\sigma}\cdot\hat{p}}{2mc}\phi$ we recall equation 2.66 which described the normalization of the two component wave function $\psi = \begin{pmatrix}\phi\\ \chi\end{pmatrix}$ we can now write:

$$\int d^3x(\phi^\dagger\phi + \chi^\dagger\chi) = \int d^3x\left(\phi^\dagger\phi + \phi^\dagger\frac{(\sigma\cdot\hat{p})^2}{4m^2c^2}\phi\right)$$

and we can conclude that we can relate the normalization of ϕ and ψ by writing:

$$\psi = \left(1+\frac{\hat{p}^2}{4m^2c^2}\right)^{\frac{1}{2}}\phi \approx \left(1+\frac{\hat{p}^2}{8m^2c^2}\right)\phi$$

which gives us the relation:

$$\phi \approx \left(1 - \frac{\hat{p}^2}{8m^2c^2}\right)\psi \tag{2.75}$$

where ψ, is the solution to the Schrödinger equations in 2.74 correctly normalized. We will now substitute for ϕ in equation (2.74). We now obtain the following eigenvalue equation for E, the energy which now is taken to contain all first order relativistic corrections:

$$\begin{aligned} E\left(1 - \frac{\hat{p}^2}{8m^2c^2}\right)\psi &= eV\left(1 - \frac{\hat{p}^2}{8m^2c^2}\right)\psi \\ &+ \frac{(\vec{\sigma}\cdot\hat{p})}{(2m)}\frac{1}{\left(1 + \frac{E}{2mc^2} - \frac{eV}{2mc^2}\right)}(\vec{\sigma}\cdot\hat{p})\left(1 - \frac{\hat{p}^2}{8m^2c^2}\right)\psi \end{aligned} \tag{2.76}$$

We now expand the term $\left(1 + \frac{E}{2mc^2} - \frac{eV}{2mc^2}\right)^{-1}$ to first order in mc^2 and re-arrange some of the terms, recalling that $(\vec{\sigma}\cdot\hat{p})^2 = \hat{p}^2$ we obtain:

$$\begin{aligned} E\psi &= (E - eV)\frac{\hat{p}^2}{8m^2c^2}\psi + eV\psi \\ &+ \left(\frac{\hat{p}^2}{2m} - \frac{\hat{p}^4}{16m^3c^2} - \frac{(\vec{\sigma}\cdot\hat{p})E\vec{\sigma}\cdot\hat{p}}{4m^2c^2} + \frac{(\vec{\sigma}\cdot\hat{p})eV(\vec{\sigma}\cdot\hat{p})}{4m^2c^2}\right)\psi \end{aligned} \tag{2.77}$$

where terms with powers of c higher than 2 have been ignored. We proceed by noting that E on the right hand side is a number, but in order to explicitly separate the relativistic corrections we can write $E\psi = \left(\frac{\hat{p}^2}{2m} + eV\right)\psi$ as an approximation. Thus we move E right next to ψ and then replace it by an operator valued approximation. Note the eV does not commute with $\hat{p}$ or $\hat{p}^2$ so taking careful account of the order of appearance of $\hat{p}$ we obtain:

$$E\psi = \frac{\hat{p}^4}{16m^3c^2}\psi + \frac{\hat{p}^2}{8m^2c^2}eV\psi = eV\frac{\hat{p}^2}{8m^2c^2}\psi + eV\psi \tag{2.78}$$

$$+\left(\frac{\hat{p}^2}{2m} - \frac{\hat{p}^4}{16m^3c^2} - \frac{\hat{p}^4}{8m^3c^2} - \frac{\hat{p}^2}{4m^2c^2}eV + \frac{(\vec{\sigma}\cdot\hat{p})eV(\vec{\sigma}\cdot\hat{p})}{4m^2c^2}\right)\psi$$

Simplifying somewhat we obtain:

$$\begin{aligned} E\psi &= eV\psi + \frac{\hat{p}^2}{2m}\psi - \frac{\hat{p}^4}{8m^3c^2}\psi - \frac{e}{8m^2c^2}(\hat{p}^2V + V\hat{p}^2)\psi \\ &+ \frac{(\vec{\sigma}\cdot\hat{p})eV(\vec{\sigma}\cdot\hat{p})}{4m^2c^2}\psi \end{aligned} \tag{2.79}$$

It is heartening to see the terms of $eV\psi$ and $\frac{\hat{p}^2}{2m}\psi$ on the right hand side showing that we have a Schrödinger like equation with some corrections. We now proceed to simplify the term:

$$\frac{-e}{8m^2c^2}(\hat{p}^2V + V\hat{p}^2)\psi + \frac{(\vec{\sigma}\cdot\hat{p})eV(\vec{\sigma}\cdot\hat{p})}{4m^2c^2}\psi \tag{2.80}$$

We re-write equation (2.80) in a slightly modified form showing the components explicitly :

$$\frac{e}{8m^2c^2}[2\sigma_i\hat{p}_iV\sigma_j\hat{p}_j - \hat{p}_i\hat{p}_iV - V\hat{p}_i\hat{p}_i]\psi \tag{2.81}$$

Recall that $\hat{p}_i = -i\hbar\partial_i$ (or $-i\hbar\vec{\nabla}$) and that $\sigma_i\sigma_j = \delta_{ij} + i\ell_{ijk}\sigma_k$. Using these definitions and operating to the right we can write down:

$$\frac{e}{8m^2c^2}((2\delta_{ij} + 2i\ell_{ijk})(Vp_ip_j\psi - i\hbar\partial_iVp_j\psi)$$

$$+\hbar^2(\partial_i\partial_iV)\psi - i\hbar\partial_iVp_i\psi - i\hbar\partial_iVp_i\psi - 2Vp_ip_i\psi)$$

Reinserting ∇^2 for $\partial_i\partial_i$ and dropping $\vec{p}\times\vec{p}$ we obtain:

$$(\frac{e\hbar^2\nabla^2V}{8m^2c^2} - \frac{e\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{p}\times\vec{\nabla}V))\psi \tag{2.82}$$

Note that $\vec{\nabla}V = \frac{\vec{x}}{|\vec{x}|}\frac{\partial V}{\partial|\vec{x}|}$ and $\vec{p}\times\vec{\nabla}V$ is $\vec{p}\times\frac{\vec{x}}{|\vec{x}|}\frac{\partial V}{\partial|\vec{x}|}$ and that $\vec{p}\times\vec{x} = -\vec{L}$ where $\vec{L}$ is the orbital angular momentum. Putting all of this in and using $|\vec{x}| = r$ we have:

$$(\frac{e\hbar^2\nabla^2V}{8m^2c^2} + \frac{e\hbar}{4m^2c^2}\vec{\sigma}\cdot\vec{L}\frac{1}{r}\frac{dV}{dr})\psi \tag{2.83}$$

Identifying $\frac{\hbar}{2}\vec{\sigma}$ as $\vec{S}$ and $\nabla^2V = Ze\delta(r)$ as the charge density of a point like nucleus (at origin $r = 0$) and collecting all terms we rewrite equation (2.79)

$$E\psi = \frac{\hat{p}^2}{2m}\psi - eV\psi - \frac{\hat{p}^4}{8m^3c^2}\psi - \frac{Ze^2\hbar^28\delta(r)}{8m^2c^2}\psi - \frac{e}{2m^2c^2}\vec{S}\cdot\vec{L}\frac{1}{r}\frac{\partial V}{\partial r}\psi \tag{2.84}$$

where we have set the electron charge to $e = -|e|$. Note the first two terms on the right hand side of equation (2.84) are the unperturbed Hamiltonian, the third the relativistic correction to the kinetic energy of the orbiting electron, the fourth the Darwin term that switches on for S-waves only and the fifth the spin-orbit coupling. All these terms are inserted by hand as perturbations when using the Schrödinger equation to solve for the hydrogen atom. As we have seen they are contained in the Dirac equation. The Lamb shift is an additional correction which emerges from the second quantized theory and is beyond the scope of this book.

In the preceding analysis we have used a rather simple way to decouple the relativistic and non-relativistic components of the Dirac equation. By using the observation that components that are large or small are coupled by the $\left(\begin{array}{c|c} 0 & \vec{\sigma} \\ \hline \vec{\sigma} & 0 \end{array}\right) = \vec{\alpha}$ terms and the β terms $(\beta eV, \beta mc^2)$ leave the ϕ and χ uncoupled, an elegant and systematic procedure – the Fouldy-Wolthuysen transformation can be used to decouple the large and small components of the Dirac equation's solutions. Several excellent texts describe this procedure in detail [8] [7] [10]. To summarize we see that the Dirac equation intrinsically contains all the information to include spin $\frac{1}{2}$ particles and the entire gamut of fine structure corrections. In addition to this Dirac's conjecture of unoccupied negative energy states being equivalent to positively charged electrons was also confirmed in the observation of the positron. These together more than justify the viability of the Dirac equation for describing spin $\frac{1}{2}$ particles in an electromagnetic field with the introduction of minimal coupling.

Chapter 3

Lorentz covariance of the Dirac equation and plane wave solutions

In this chapter we will first prove the relativistic covariance of the Dirac equation *i.e.* that the *form* of the Dirac equation does not change from one Lorentz frame to another. We'll begin by writing the Dirac equation in a form where the space and time co-ordinates are treated on an equal footing introducing the γ matrices which will take the place of the α and β matrices. Then a Lorentz transformation will be defined for the wave function itself and we'll verify that this transformation satisfies certain conditions. We will do this for an infinitesimal Lorentz transformation as defined in chapter 1 section 1.8. Once we know this rule for transforming the wave function from one Lorentz frame to another we will use it to obtain plane wave solutions to the Dirac equation by applying an infinite number of infinitesimal boosts to obtain a finite boost that allows us to view it from a frame in which the wave propagates in the positive sense (as per convention) along an arbitrary direction. We will then write down these solutions and define their orthogonality relations and will show from the properties under rotation of the Dirac equation solutions that they are indeed spinors.

We will also classify the transformation properties of certain quantities-bilinear covariants-constructed from ψ, $\psi^\dagger$ and the γ matrices mentioned above. We will also demonstrate discrete symmetries of the Dirac equation under the charge conjugation, parity and time-reversal operations. These discrete operations are crucial in the interpretation of negative energy electrons travelling backward in time as being positrons travelling forward in time which is how positrons are treated in the Feynman propagator approach to quantum electrodynamics.

We'll also write down the spin projection and energy projection operators and define the completeness relations of the solutions. Finally we'll show that the solutions are eigenstates of helicity and will define chirality.

3.1 Covariant representation of the Dirac Equation

We have so far written the Dirac equation with the time co-ordinate given a special place which as we know is called the Hamiltonian form of the Dirac equation. Note the Schrödinger equation can be written in this for too. We've so far used the Dirac equation in the form

$$i\hbar\frac{\partial}{\partial t}\psi = (-ic\hbar\vec{\alpha}\cdot\vec{\nabla} + V(|\vec{x}|) + \beta mc^2)\psi \tag{3.1}$$

In order to treat space and time on an equal footing we multiply the Dirac equation equation 3.1 on the right by the matrix β

$$\beta = \gamma^0 \tag{3.2}$$

and the products of β with each α matrix $\beta\alpha_k$ by γ^k:

$$\gamma^k = \beta\alpha_k \tag{3.3}$$

recalling that $\beta^2 = (\gamma^0)^2 = \mathbb{I}$, we can write down

$$i\hbar\gamma^0\frac{\partial\psi}{\partial t} = -i\hbar c(\gamma^1\frac{\partial\psi}{\partial x^1} + \gamma^2\frac{\partial\psi}{\partial x^2} + \gamma^3\frac{\partial\psi}{\partial x^3}) + mc^2\psi \tag{3.4}$$

Rearranging and dividing by c we obtain the covariant (although we have not yet actually proven that it is relativistically covariant) form of the Dirac equation :

$$i\hbar\left(\gamma^0\frac{\partial\psi}{\partial x^0} + \gamma^1\frac{\partial\psi}{\partial x^1} + \gamma^2\frac{\partial\psi}{\partial x^2} + \gamma^3\frac{\partial\psi}{\partial x^3}\right) - mc\psi = 0 \tag{3.5}$$

This is the final form of the Dirac equation that we will use, the only other modification will be to change the units to natural units described in section 1.15. The reader can verify easily that with the definitions 3.2-3.3 of the γ matrices in terms of the α and β matrices (equations 2.18-2.19):

$$\beta = \gamma^0 = \left(\begin{array}{c|c}\mathbb{I} & 0\\ \hline 0 & -\mathbb{I}\end{array}\right) \tag{3.6}$$

$$(\vec{\gamma})^k = \gamma^k = \gamma_0\alpha^k = \left(\begin{array}{c|c}0 & \sigma_k\\ \hline -\sigma_k & 0\end{array}\right) \tag{3.7}$$

where each block matrix is a 2×2 matrix embedded in a 4×4 matrix where the σ_k, are the 2×2 Pauli matrices.

We write these down explicitly once as the explicit form may be needed for quick reference in further chapters.

$$\gamma_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \gamma_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \gamma_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{3.8}$$

The γ^0 and γ^1, γ^2, γ^3 are collectively referred to as the γ matrices and are denoted by γ^μ. When we want to refer to the matrices associated with time and space indices separately we use $(\gamma^0, \vec{\gamma})$ or $(\gamma^0,\ \gamma^j)$ as notation. The reader is asked to note that despite the 4-vector notation these do not transform as a 4-vector-we in fact assume them to be the same from frame to frame. However they are associated with the four space-time co-ordinates as in equation 2.24. Note that all the γ matrices are not Hermitian-only γ^0 is, the reader is encouraged to check this. We now revisit their anti-commutation properties:

1. Note that, trivially $(\gamma^0)^2 = \mathbb{I}$ but that $(\gamma^k)^2 = -\mathbb{I}$, here the k refers to the spatial indices only.
2. We begin with the anti-commutator: $\{\gamma^i, \gamma^j\} = \{\alpha_i\beta, \alpha_j\beta\} = \alpha_i\beta\alpha_j\beta + \alpha_j\beta\alpha_i\beta$. Now recall $\{\alpha_i, \beta\} = 0$ thus $\alpha_i\beta = -\beta\alpha_i$ and we can use this to pull all the βs next to each other: (remembering to pick up a (-) sign each time):

$$\begin{aligned}\alpha_i\beta\alpha_j\beta + \alpha_j\beta\alpha_i\beta &= -\alpha_i\beta\beta\alpha_j - \alpha_j\beta\beta\alpha_i = -\alpha_i\beta^2\alpha_j - \alpha_j\beta^2\alpha_i \\ &= -(\alpha_i\alpha_j + \alpha_j\alpha_i)\end{aligned}$$

but we already know that $\{\alpha_i, \alpha_j\} = 2\delta^{ij}\mathbb{I}$ thus:

$$\{\gamma^i, \gamma^j\} = -2\delta ij\mathbb{I} \tag{3.9}$$

where we have so far dealt with only the γs that carry a spatial index. Now since $\beta^2 = \gamma^{02} = \mathbb{I}$ thus we can summarize the anti commutation relations such as:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{I} \tag{3.10}$$

where $\mathbb{I}$ is the 4×4 identity matrix but $g^{\mu\nu}$ are the $(\mu\nu)^{th}$ elements of the matrix $g^{\mu\nu}$ and not the matrix itself.

3. It is also easy to verify the useful property $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. Note also that γ^0 is Hermitian however the γ^i or $\gamma^{1,2,3}$ are anti-Hermitian with $(\gamma^i)^\dagger = -\gamma^i$

At this point we will also introduce the Feynman "slash"notation where $\gamma^\mu B_\mu = \not{B}$ and thus rewrite equation 2.23 in a more compact form as:

$$(i\hbar\not{\partial} - mc)\psi = 0 \tag{3.11}$$

with:

$$\gamma^\mu \frac{\partial}{\partial x^\mu} = \gamma^\mu \partial_\mu = \not{\partial}$$

. The equation 3.11 is the form of the Dirac equation that we will use. In order to consider the Dirac equation in the presence of an electromagnetic field we replace $i\hbar\partial_\mu$ by $i\hbar\partial_\mu - \frac{e}{c}A_\mu$ and so we write

$$(i\hbar\not{\partial} - \frac{e}{c}\not{A} - mc)\psi = 0 \tag{3.12}$$

with the Feynman-slash now used for denoting $\gamma^\mu A_\mu(x)$. Here e represents the charge of the electron which is $e = - \mid e \mid$ note that we will discuss the interpretation of the negative energy solutions as positrons later in this chapter where the positron charge will be $\mid e \mid$.

Finally the reader is asked to recall the probability current in equation 2.29, and to verify that in terms of the gamma matrices we can simply write:

$$j^\mu = (\rho, \vec{\jmath}) = (\psi^\dagger\psi, \psi^\dagger\gamma^0\vec{\gamma}\psi) = (\psi^\dagger\gamma^0\gamma^\mu\psi) = \bar{\psi}\gamma^\mu\psi \tag{3.13}$$

Here we have defined the adjoint wave-function $\bar{\psi}$ in equation 3.13, since $\psi^\dagger\gamma^0$ occurs so often we have given it its' own symbol *i.e.* $\bar{\psi}$-the adjoint spinor.

$$\bar{\psi} = \psi^\dagger\gamma^0 \tag{3.14}$$

3.2 The boost matrix S for Lorentz transforming the 4-component spinor wave function $\psi(x)$

In order to show that the Dirac Equation retains its form under Lorentz transformation (i.e. show that it is Lorentz covariant) we have to write down the Dirac equation in an boosted and an unboosted frame and show that the form is the same. We know how to transform x^μ and ∂^μ but not the 4-component wave function ψ. We therefore postulate the existence of a 4×4 matrix S relating $\psi(x)$ and $\psi'(x')$ in the following manner:

$$\psi'(x')_\theta = S_{\theta\alpha}\psi(x)_\alpha$$

with the indices suppressed for convenience we will simply write:

$$\psi'(x') = S\psi(x) \tag{3.15}$$

Let's now write down the Dirac equation in both the boosted and unboosted frames. We have to remember that whilst S transforms ψ the derivative operator transforms like a 4-vector. In the unboosted frame it is simply

$$i\hbar\gamma^\mu\partial_\mu\psi(x) - mc\psi(x) = 0 \tag{3.16}$$

and in the boosted frame we should have accordingly (choosing a different index for the sum)

$$(i\hbar\gamma^\nu\partial'_\nu - mc)\psi'(x') = 0 \tag{3.17}$$

Note that here we have assumed that the γ matrices remain the same in the transformed frame. This is rigorously justified in several texts [8 10 12 7]. Explicitly inserting $S\psi(x)$ in place of $\psi(x')$ with indices suppressed in Eqn. 3.17 according to the definition in Eqn. 3.15 we obtain:

$$i\hbar\gamma^\nu\partial'_\nu S\psi(x) - mcS\psi(x) = 0 \tag{3.18}$$

We are going to compare a modified form of Eqn. 3.18 with the unboosted equation 3.16. We will do this by expressing ∂_μ in Eqn. 3.16 in terms of ∂'_μ using the property of transformation of covariant vectors (please see chapter 1):

$$\partial'_\nu = \frac{\partial x^\mu}{\partial x^{\nu\prime}}\frac{\partial}{\partial x^\mu} = (\Lambda^{-1})^\mu{}_\nu\partial_\mu \tag{3.19}$$

where $\Lambda^\nu{}_\mu = \frac{\partial x'^\nu}{\partial x^\mu}$ is the usual Lorentz transform matrix for 4-vectors. Note the elements of Λ^{-1} define the Lorentz transformation of an covariant vector as specified in equation 1.26-simply the inverse of the Lorentz transformation. We now rewrite 3.17 using 3.19

$$(i\hbar\gamma^{\nu}(\Lambda^{-1})^{\mu}{}_{\nu}\partial_{\mu} - mc)\psi'(x') = 0 \tag{3.20}$$

and now using $S\psi$ for ψ' suppressing arguments of x' etc we have

$$(i\hbar\gamma^{\nu}(\Lambda^{-1})^{\mu}{}_{\nu}\partial_{\mu} - mc)S\psi = 0 \tag{3.21}$$

multiplying on the left with S^{-1} we obtain:

$$(i\hbar S^{-1}\gamma^{\nu}S(\Lambda^{-1})^{\mu}{}_{\nu}\partial_{\mu} - mc)\psi = 0 \tag{3.22}$$

we now compare equation 3.22 with equation 3.16 and note that both are now equations for ψ, suffice to say if we can find an S such that

$$S^{-1}\gamma^{\nu}S(\Lambda^{-1})^{\mu}{}_{\nu} = \gamma^{\mu} \tag{3.23}$$

can be satisfied then we have a rule for transforming ψ from frame to frame and will have prooved the covariance of the Dirac equation. Equation 3.23 can be modified by multiplying both sides by the Lorentz transform matrix to obtain a slightly less cumbersome form of this equation:

$$S^{-1}\gamma^{\nu}S = \gamma^{\mu}\Lambda^{\nu}{}_{\mu} \tag{3.24}$$

For simplicity's sake we'll pick $\Lambda^{\nu}{}_{\mu}$ to be an infinitesimal Lorentz boost, it will logically follow that if an S can be found to satisfy Eqn. 3.24 for an infinitesimal boost, then one can be found for a finite boost. We will simply verify that a particular choice for S, satisfies the relation 3.24 for the case of an infinitesimal Lorentz boost. The reader is referred to section 1.8 of the first chapter for a review of infinitesimal transforms for a 4-vector.

The claim is made that the following choice for S satisfies this relation:

$$S = \mathbb{I} - \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta} \tag{3.25}$$

where $\sigma_{\alpha\beta} = \frac{i}{2}[\gamma_{\alpha}\gamma_{\beta} - \gamma_{\beta}\gamma_{\alpha}]$ (thus e.g. $\sigma_{12} = \frac{i}{2}[\gamma_1\gamma_2 - \gamma_2\gamma_1]$ etc) and the $\epsilon^{\alpha\beta}$ were defined in Eqns. 1.45-1.51 is an appropriate choice for S and will satisfy the relation 3.24, it should be clear $\epsilon^{\alpha\beta}$ will denote that $(\alpha\beta)^{th}$ element of the appropriate (x, y or z) boost matrix. Now $\alpha\beta$ are space time indices operating on 4-vectors like A_{α}, ∂_{β} etc and the upper or lower indices of the γ matrices on the beta in γ_{β} but not of course on the indices that operate on the components of ψ thus not on $\sigma\rho$ in the product $\gamma_{\beta,\sigma\rho}\psi_{\rho}$

We now move on to the verification of the relation 3.24 using S as defined in 3.25. Before doing so we need remind ourselves of the following:

1. The inverse of S can be easily verified :
$$S^{-1} = \mathbb{I} + \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta} \tag{3.26}$$
by doing performing the operation $S^{-1}S$ and keeping terms to first order in ϵ.
2. As we shall see we will need an identity involving lowered index and raised index γ matrices. Without much ado we derive it here. The anticommutation relation $\{\gamma^{\mu}, \gamma^{\nu}\} = 2g_{\mu\nu}$ gives us
$$\gamma^{\mu}\gamma^{\nu} = 2g^{\mu\nu} - \gamma^{\nu}\gamma^{\mu} \tag{3.27}$$
Multiplying on Eqn. 3.27 on the left with $g_{\alpha\mu}$ we obtain :
$$\gamma_{\alpha}\gamma^{\nu} = 2g^{\nu}{}_{\alpha} - \gamma^{\nu}\gamma_{\alpha} \tag{3.28}$$
We will use this several times in the developments that follow. Note that $g^{\nu}{}_{\alpha}$ is named as such because of the index lowering operation with $g_{\alpha\mu}$ was performed, it is the same object as $\delta^{\nu}{}_{\alpha}$ (trivially) and the reader is reminded $\delta^{\nu}{}_{\alpha}$ is equals zero if the two indices are not the same and 1 if they are, same for $g^{\nu}{}_{\alpha}$.

So lets now write out

$$S^{-1}\gamma^{\nu}S = (\mathbb{I} + \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta})\gamma^{\nu}(\mathbb{I} - i\frac{i}{4}\sigma_{\sigma\rho}\epsilon^{\sigma\rho}) \tag{3.29}$$

Using $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_{\mu}, \gamma_{\nu}]$ in Eqn. 3.29, and multiplying out factors of i, $\frac{1}{2}$, $\frac{1}{4}$ we obtain:

$$S^{-1}\gamma^{\nu}S = (\mathbb{I} - \frac{1}{8}[\gamma_{\alpha},\, \gamma_{\beta}]\epsilon^{\alpha\beta})\gamma^{\nu}(\mathbb{I} + \frac{1}{8}[\gamma_{\sigma},\, \gamma_{\rho}]\epsilon^{\sigma\rho}) \tag{3.30}$$

expanding Eqn. 3.30 and keeping terms up to first order in $\epsilon^{\alpha\beta}$ we have:

$$\gamma^{\nu} - \frac{1}{8}[\gamma_{\alpha},\, \gamma_{\beta}]\epsilon^{\alpha\beta}\gamma^{\nu} + \frac{\gamma^{\nu}}{8}[\gamma_{\sigma},\, \gamma_{\rho}]\epsilon^{\sigma\rho} \tag{3.31}$$

next we expand out the commutators in 3.31 and obtain:

$$\gamma^{\nu} - \frac{1}{8}\gamma_{\alpha}\gamma_{\beta}\gamma^{\nu}\epsilon^{\alpha\beta} + \frac{1}{8}\gamma_{\beta}\gamma_{\alpha}\gamma^{\nu}\epsilon^{\alpha\beta} + \frac{1}{8}\gamma^{\nu}\gamma_{\sigma}\gamma_{\rho}\epsilon^{\sigma\rho} - \frac{1}{8}\gamma^{\nu}\gamma_{\rho}\gamma_{\sigma}\epsilon^{\sigma\rho} \tag{3.32}$$

Note that $\alpha\beta$ and $\sigma\rho$ are summed over and it is alright for simplicity's sake to just use $\alpha\beta$ and doing this we obtain:

$$\gamma^\nu - \frac{1}{8}(\gamma_\alpha\gamma_\beta\gamma^\nu - \gamma_\beta\gamma_\alpha\gamma^\nu - \gamma^\nu\gamma_\alpha\gamma_\beta + \gamma^\nu\gamma_\beta\gamma_\alpha)\epsilon^{\alpha\beta} \tag{3.33}$$

Now recall Eqn. 3.28, using $\gamma^\nu\gamma_\alpha = 2g^\nu{}_\alpha - \gamma_\alpha\gamma^\nu$ and $\gamma^\nu\gamma_\beta = 2g^\nu{}_\beta - \gamma_\alpha\gamma^\nu$ we will move γ^ν to the right of third and fourth terms in the brackets in Eqn. 3.33. Doing this once we obtain:

$$-\gamma^\nu\gamma_\alpha\gamma_\beta + \gamma^\nu\gamma_\beta\gamma_\alpha = -(2g^\nu{}_\alpha\gamma_\beta - \gamma_\alpha\gamma^\nu\gamma_\beta) + (2g^\nu{}_\beta\gamma_\alpha + \gamma_\beta\gamma^\nu\gamma_\alpha) \tag{3.34}$$

We once again move γ^ν to the right in the terms that contain it in Eqn. 3.34 and obtain:

$$\begin{aligned} -\gamma^\nu\gamma_\alpha\gamma_\beta + \gamma^\nu\gamma_\beta\gamma_\alpha = &-(2g^\nu{}_\alpha\gamma_\beta - 2g^\nu{}_\beta\gamma_\alpha + \gamma_\alpha\gamma_\beta\gamma^\nu) \\ &+(2g^\nu{}_\beta\gamma_\alpha - 2.g^\nu{}_\alpha\gamma_\beta + \gamma_\beta\gamma_\alpha\gamma^\nu) \end{aligned} \tag{3.35}$$

Substituting for $-\gamma^\nu\gamma_\alpha\gamma_\beta + \gamma^\nu\gamma_\beta\gamma_\alpha$ in Eqn. 3.33 we obtain

$$\begin{aligned} \gamma^\nu - \frac{1}{8}(\gamma_\alpha\gamma_\beta\gamma^\nu - \gamma_\beta\gamma_\alpha\gamma^\nu &- (2g^\nu{}_\alpha\gamma_\beta - 2g^\nu{}_\beta\gamma_\alpha + \gamma_\alpha\gamma_\beta\gamma^\nu) \\ &+(2g^\nu{}_\beta\gamma_\alpha - 2g^\nu{}_\alpha\gamma_\beta + \gamma_\beta\gamma_\alpha\gamma^\nu))\epsilon^{\alpha\beta} \end{aligned} \tag{3.36}$$

and we can see simply that all the terms containing γ^ν cancel and we are left with:

$$\gamma^\nu - \frac{1}{8}(-2g^\nu{}_\alpha\gamma_\beta + 2g^\nu{}_\beta\gamma_\alpha + 2g^\nu{}_\beta\gamma_\alpha - 2g^\nu{}_\alpha\gamma_\beta)\epsilon^{\alpha\beta} \tag{3.37}$$

Since the g^α_βs are zero unless their indices are the same we can sum over one of the indices of the $\epsilon^{\alpha\beta}$ to obtain:

$$\gamma^\nu - \frac{1}{8}[-2\gamma_\beta\epsilon^{\nu\beta} + 2\gamma_\alpha\epsilon^{\alpha\nu} + 2\gamma_\alpha\epsilon^{\alpha\nu} - 2\gamma_\beta\epsilon^{\nu\beta}] \tag{3.38}$$

All the repeated indices are summed over, we can therefore replace each repeated index with one index thus we pick μ doing this we obtain trivially:

$$\gamma^\nu - \frac{1}{8}(-2\epsilon^{\nu\mu}\gamma_\mu + 2\epsilon^{\mu\nu}\gamma_\mu + 2\epsilon^{\mu\nu}\gamma_\mu - 2\epsilon^{\nu\mu}\gamma_\mu) \tag{3.39}$$

Using Eqn. 1.51 or $\epsilon^{\nu\mu} = -\epsilon^{\mu\nu}$ etc we obtain:

$$\gamma^\nu + \gamma_\mu\epsilon^{\mu\nu} = \gamma^\mu(\delta^\nu{}_\mu + \epsilon^\nu{}_\mu) \tag{3.40}$$

Recall that we defined ${\delta^\nu}_\mu + {\epsilon^\nu}_\mu = {\Lambda^\nu}_\mu$ in Eqn 1.40:

Thus $S^{-1}\gamma^\nu S = \gamma^\mu({\delta^\nu}_\mu + {\epsilon^\nu}_\mu) = \gamma^\mu {\Lambda^\nu}_\mu$ recalling that this is what we had set out to validate. Thus S is an appropriate choice for transforming the spinor $\psi(x)$ in the Dirac equation and the Dirac equation is Lorentz-covariant and also that

$$S^{-1}\gamma^\nu S = \gamma^\mu {\Lambda^\nu}_\mu \tag{3.41}$$

We will use Eqn. 3.41 repeatedly in the next section.

3.3 Transforming $\psi^\dagger$, defining $S^\dagger$

We need to consider how $\psi^\dagger$ transforms under a Lorentz transformation, it should be clear that this should transform as $\psi^{\dagger\prime} = \psi^\dagger S^\dagger$. This will be needed in the next section when we form objects with various transformation properties by "sandwiching" gamma matrices between $\psi^\dagger$ and ψ. For this of course we need to work out $S^\dagger$, so this section is a mathematical piece dedicated to this small task.

$$S = \mathbb{I} - \frac{i}{4}\frac{i}{2}[\gamma_\mu, \gamma_\nu]\epsilon^{\mu\nu} \text{ with } \sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \tag{3.42}$$

we will need to obtain the Hermitian conjugate of $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$ which of course involves the Hermitian conjugate (transpose followed by complex conjugation) of a product of matrices. We will do this for a product of two other matrices $\mathbb{M}_1$ and $\mathbb{M}_2$ which we will denote by $\mathbb{M}_3$. We use the superscript ^{T} for the transpose of a matrix , and the subscripts $_I$ and $_R$ for imaginary and real parts of a matrix:

$$\mathbb{M}_{3,\sigma\rho} = \mathbb{M}_{1,\sigma\alpha}\mathbb{M}_{2,\alpha\rho}$$

then the elements of the transpose of $\mathbb{M}_3$ in terms of the elements of the transposes of $\mathbb{M}_1$ and $\mathbb{M}_2$ are:

$$(\mathbb{M}_3)^T_{\rho\sigma} = \mathbb{M}_{3,\sigma\rho} = (\mathbb{M}^T)_{2,\sigma\alpha}(\mathbb{M}^T)_{1,\alpha\rho}$$

where the order of the matrix multiplication is apparent in the ordering of the indices. We can see that in general $(\mathbb{M}_1\mathbb{M}_2)^T = \mathbb{M}_2^T\mathbb{M}_1^T$. Now assuming these matrices are complex we split them up into their real and imaginary parts:

$$\mathbb{M}_1\mathbb{M}_2 = (\mathbb{M}_{1,R}\mathbb{M}_{2,R} - \mathbb{M}_{1,I}\mathbb{M}_{2,I}) + i(\mathbb{M}_{1,R}\mathbb{M}_{2,I} + \mathbb{M}_{1,I}\mathbb{M}_{2,R}) \quad (3.43)$$

Note that applying the rule for the transpose of a product of two matrices and the definition of the Hermitian conjugate we obtain:

$$\begin{aligned}(\mathbb{M}_1\mathbb{M}_2)^\dagger = ((\mathbb{M}_1\mathbb{M}_2)^T)^* = (\mathbb{M}_2^T\mathbb{M}_1^T)^* = \\ (\mathbb{M}_{2,R}^T\mathbb{M}_{1,R}^T - \mathbb{M}_{2,I}^T\mathbb{M}_{1,I}^T) - i(\mathbb{M}_{2,R}^T\mathbb{M}_{1,I}^T + \mathbb{M}_{2,I}^T\mathbb{M}_{1,R}^T)\end{aligned} \quad (3.44)$$

its easy to see by inspecting Eqns 3.44 and using the definition of a Hermitian conjugate that:

$$(\mathbb{M}_1\mathbb{M}_2)^\dagger = \mathbb{M}_2^\dagger\mathbb{M}_1^\dagger \quad (3.45)$$

Now we turn back to determining $S^\dagger$. Recall that $S = \mathbb{I} - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}$ where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$. Taking the Hermitian conjugate explicitly we obtain

$$S^\dagger = \mathbb{I} + \frac{i}{4}(\sigma_{\mu\nu})^\dagger\epsilon^{\mu\nu} \quad (3.46)$$

It is clear that

$$\sigma_{\mu\nu}^\dagger = (\frac{i}{2}[\gamma_\mu, \gamma_\nu])^\dagger = \frac{-i}{2}[(\gamma_\mu\gamma_\nu)^\dagger - (\gamma_\nu\gamma_\mu)^\dagger] \quad (3.47)$$

using Eqn. 3.45 this is trivially rewritten

$$\sigma_{\mu\nu}^\dagger = \frac{-i}{2}[\gamma_\nu^\dagger\gamma_\mu^\dagger - \gamma_\mu^\dagger\gamma_\nu^\dagger] \quad (3.48)$$

using $\gamma_\mu^\dagger = \gamma_0\gamma_\mu\gamma_0$ we obtain

$$\sigma_{\mu\nu}^\dagger = \frac{-i}{2}[\gamma_0\gamma_\nu\gamma_0\gamma_0\gamma_\mu\gamma_0 - \gamma_0\gamma_\mu\gamma_0\gamma_0\gamma_\nu\gamma_0] \quad (3.49)$$

using $(\gamma_0)^2 = \mathbb{I}$ this is easily seen to be

$$\sigma_{\mu\nu}^\dagger = \frac{-i}{2}\gamma_0[\gamma_\nu\gamma_\mu - \gamma_\mu\gamma_\nu]\gamma_0 = \frac{i}{2}\gamma_0[\gamma_\mu, \gamma_\nu]\gamma_0 = \gamma_0\sigma_{\mu\nu}\gamma_0 \quad (3.50)$$

where reversing the order of the indices in the commutator removes one minus sign.

Using Eqn. 3.50 and $(\gamma_0)^2$ to substitute for $\sigma_{\mu\nu}$ and $\mathbb{I}$ in Eqn. 3.46 we obtain

$$S^\dagger = \mathbb{I} + \frac{i}{4}(\sigma_{\mu\nu})^\dagger \epsilon^{\mu\nu} = \gamma_0\gamma_0 + \frac{i}{4}\gamma_0\sigma_{\mu\nu}\gamma_0\epsilon^{\mu\nu} \tag{3.51}$$

If we now recall the defnition of S^{-1} in Eqn. 3.26 we can see easily that

$$S^\dagger = \gamma_0 S^{-1}\gamma_0 \tag{3.52}$$

or $S^\dagger = \gamma^0 S^{-1}\gamma^0$.

We will now move on to form several quantities involving $\bar{\psi} = \psi^\dagger\gamma_0\ \psi$ and the γ matrices. We will also document their transformation properties, for this it should be clear that we will need to use $S^\dagger$ and S extensively. Examples of second rank tensors, pseudo vectors and scalars will also be given and verified.

3.4 Selected Bilinear Covariants

We will now examine the transformation properties of selected bilinear covariants constructed from $\psi^\dagger$, ψ and the γ matrices. For example in the previous chapter we had without proof taken $\bar{\psi}\gamma^\mu$ to be a contravariant 4-vector, in this section we will actually demonstrate that rigorously.

1. Let us first consider the $\bar{\psi}\psi = \psi^\dagger\gamma^0\psi$ where the adjoint spinor $\bar{\psi}$ was first defined in Eqn. 3.14. We can use our rule for transforming the ψs and it is easy to see that under a Lorentz transform this changes to $\psi^\dagger S^\dagger\gamma^0 S\psi$:

$$\bar{\psi}'\psi' = \psi^\dagger S^\dagger\gamma^0 S\psi = \psi^\dagger\gamma^0 S^{-1}\gamma^0\gamma^0 S\psi = \psi^\dagger\gamma^0\psi \tag{3.53}$$

   we conclude therefore that the quantity $\bar{\psi}\psi = \psi^\dagger\gamma^0\psi$ is a Lorentz scalar.
2. We begin by reminding the reader of the probability current $c\psi^\dagger\gamma^0\gamma^\mu\psi = \bar{\psi}\gamma^\mu\psi$ which was derived in chapter 2 Eqn. 3.13. We will prove here that it indeed has the properties of Lorentz covariance of a 4-vector. By the properties of the matrix S for infinitesimal boosts if the current in one frame of reference is

$$j^\mu = c\bar{\psi}\gamma^\mu\psi = c\psi^\dagger\gamma^0\gamma^\mu\psi \tag{3.54}$$

   then in another Lorentz frame it is simply:

$$c\psi^\dagger S^\dagger \gamma^0 \gamma^\mu S\psi \tag{3.55}$$

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

$$c\psi^\dagger \gamma^0 S^{-1}\gamma^0\gamma^0\gamma^\mu S\psi = c\bar{\psi}S^{-1}\gamma^\mu S\psi \tag{3.56}$$

however we know that $S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu$, this is what we just proved in section 3.2 thus

$$\begin{aligned} j^{\mu\prime} &= c\psi^\dagger\gamma^0 S^{-1}\gamma^0\gamma^0 g^\mu S\psi = c\bar{\psi}S^{-1}\gamma^\mu S\psi = \Lambda^\mu{}_\nu c\bar{\psi}\gamma^\nu\psi \\ &= \Lambda^\mu{}_\nu j^\nu \end{aligned} \tag{3.57}$$

this is exactly the transformation property of a 4-vector hence we have prooved that j^μ is contravariant 4-vector. Note that multiplying the probability 4-current by the electrical charge of a spin $\frac{1}{2}$ particle will give us the charge current density of the particle, this is easy to verify. For a charged particle of charge q we can define:

$$J^{\mu\prime} = qc\bar{\psi}(x)\gamma^\mu\psi(x) \tag{3.58}$$

3. Let us now consider the object :

$$\bar{\psi}\gamma^\mu\gamma^\nu\psi = \psi^\dagger\gamma^0\gamma^\mu\gamma^\nu\psi$$

in another Lorentz frame, this is simply given by:

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

which can be rewritten as

$$\psi^\dagger\gamma^0 S^{-1}\gamma^0\gamma^0\gamma^\mu\gamma^\nu S\psi = \bar{\psi}S^{-1}\mathbb{I}\gamma^\mu\gamma^\nu S\psi$$

and is equal to

$$\bar{\psi}S^{-1}\gamma^\mu\gamma^\nu S\psi$$

by inserting $S^{-1}S$ between γ^μ and γ^ν we can use $S^{-1}\gamma^\mu S = \Lambda^\mu{}_\alpha\gamma^\alpha$

$$\psi^\dagger\gamma^0 S^\dagger\gamma^\mu\gamma^\nu S\psi = \bar{\psi}S^{-1}\gamma^\mu SS^{-1}\gamma^\nu S\psi = \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\bar{\psi}\gamma^\alpha\gamma^\beta\psi \tag{3.59}$$

which shows that $\bar{\psi}\gamma^\mu\gamma^\nu\psi(x)$ is a tensor of rank 2. It should be obvious that $\bar{\psi}\sigma_{\mu\nu}\psi$ is an antisymmetric tensor of rank 2 where $\sigma_{\mu\nu} = [\gamma^\mu, \gamma^\nu]$.

4 . We now define the matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{3.60}$$

(this notation is a relic of times when the γ were labelled from 1-4) note that it is trivial to prove that it anti-commutes with all the γ matrices and the reader is encouraged to check this. Note also that an odd permutation of any of the matrices in the product that defines it will incur a $-$ sign. The γ^5 matrix is given explicitly by:

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{3.61}$$

Note that $(\gamma^5)^2 = \mathbb{I}$.

We now consider the quantity $\bar{\psi}\gamma^5\psi$ under Lorentz transformations. This is simply by our own definition: $\bar{\psi}\gamma^5\psi = i\bar{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi$. Let us view this in a different Lorentz frame in the usual way:

$$\begin{aligned} &i\bar{\psi}S^{-1}\gamma^0\gamma^1\gamma^2\gamma^3 S\psi \\ &= i\bar{\psi}S^{-1}\gamma^0 SS^{-1}\gamma^1 SS^{-1}\gamma^2 SS^{-1}\gamma^3 S\psi \end{aligned} \tag{3.62}$$

where we have inserted a factor of SS^{-1} between the γs. Now we use: $S^{-1}\gamma^\mu S = \Lambda^\mu_{\ \nu}\gamma^\nu$ to obtain:

$$\begin{aligned} &i\bar{\psi}S^{-1}\gamma^0 SS^{-1}\gamma^1 SS^{-1}\gamma^2 SS^{-1}\gamma^3 S\psi \\ &= i\Lambda^0_\alpha\Lambda^1{}_\beta\Lambda^2{}_\mu\Lambda^3{}_\nu\bar{\psi}\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu\psi \end{aligned} \tag{3.63}$$

To simplify the expression 3.63 we arbitrarily choose a boost along x to define the elements of Λ. Note that the only non-zero elements of a Lorentz transformation for a boost along x are $\Lambda^0{}_0$, $\Lambda^0{}_1$, $\Lambda^1{}_0$, $\Lambda^1{}_1$, $\Lambda^2{}_2 = \Lambda^3{}_3$. Using all of these in the expression 3.63 and expanding out the sum we obtain

$$\begin{aligned} &i\bar{\psi}(\Lambda^0{}_0\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\gamma^0\gamma^0\gamma^2\gamma^3\psi + \Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\gamma^1\gamma^0\gamma^2\gamma^3 + \\ &\Lambda^0{}_0\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3\gamma^0\gamma^1\gamma^2\gamma^3 + \Lambda^0{}_1\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3\gamma^1\gamma^1\gamma^2\gamma^3)\psi \end{aligned} \tag{3.64}$$

the first and fourth terms cancel since $\gamma^0\gamma^0 = -\gamma^1\gamma^1 = \mathbb{I}$. The γ matrices in the remaining two terms are distinct and the second term is

$$
\begin{aligned}
&i\Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^1\gamma^0\gamma^2\gamma^3\psi = -i\Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi \\
&= -\Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^5\psi
\end{aligned}
$$

and the third term is trivially:

$$
i\Lambda^0_0\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi = \Lambda^0{}_0\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^5\psi
$$

So we are left with:

$$
\begin{aligned}
&-\Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^5\psi + \Lambda^0{}_0\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3\bar{\psi}\gamma^5\psi \\
&= (\bar{\psi}\gamma^5\psi)(\Lambda^0{}_0\Lambda^1{}_1\Lambda^2{}_2\Lambda^3{}_3 - \Lambda^0{}_1\Lambda^1{}_0\Lambda^2{}_2\Lambda^3{}_3)
\end{aligned}
\tag{3.65}
$$

On inspecting a matrix for a boost along x it is easy to see that Eqn 3.65 is simply the determinant of Λ times $\bar{\psi}\gamma^5\psi$, thus:

$$
\bar{\psi}\gamma^5\psi = \mathrm{Det}(\Lambda)\bar{\psi}\gamma^5\psi \tag{3.66}
$$

We know that the determinant of Λ is ± 1, it is positive for proper Lorentz transformations and negative for improper ones as defined in the discussion preceding and immediately after Eqns. 1.3-1.4. Improper transformations correspond to spatial inversion of the space co-ordinates as well as reversal of time direction. We conclude that $\bar{\psi}\gamma^5\psi$ has properties of a pseudoscalar *i.e.* it changes sign when the Lorentz transformation includes an inversion of the space co-ordinates.

5. A similar analysis to item (4) above will show that $\bar{\psi}\gamma^5\gamma^\mu\psi$ transforms as $\mathrm{Det}(\Lambda)\Lambda^\mu{}_\nu\bar{\psi}\gamma^5\gamma^\nu\psi$, which is the pseudovector property.

3.5 Discrete Symmetries of the Dirac equation: Parity, Charge Conjgation and Time Reversal

We will now consider certain discrete symmetries and transformations of the Dirac equation. We do this to prove the statement that "A negative energy solution to the Dirac equation *travelling backward in time* is equivalent to a positive energy solution of opposite charge travelling forward in time." This statement will form the basis for our treatment of positrons later in this manuscript.

1. We first consider the parity transformation which means that we will view the Dirac equation and it's solutions with the spatial co-ordinates reversed. We first discussed inversions in the context of

improper Lorentz transformations Eqns. 1.3-1.4. We will denote the space inversion or parity operator for four vectors by: $\Lambda^{\nu}_{\mu}(P)$, since it is a Lorentz transform (but an improper one). Note that the action of this operator on any 4-vector is to reverse the sign of all spatial components, thus $\Lambda^0_0(P) = 1$, $\Lambda^1_1(P) = \Lambda^2_2(P) = \Lambda^3_3(P) = -1$ all other $\Lambda^{\nu}_{\mu}(P) = 0$. Recall according to the condition Eqn. 3.41 to obtain:

$$\Lambda^{\nu}_{\mu}(P)\gamma^{\mu} = S^{-1}(P)\gamma^{\nu}S(P) \tag{3.67}$$

where the argument P of S denotes the parity transformation hence the use of $S(P)$. Using the previously defined elements of $\Lambda^{\nu}_{\mu}(P)$ we can write down:

$$\begin{aligned}
+\gamma^0 &= S^{-1}(P)\gamma^0 S(P)\\
-\gamma^1 &= S^{-1}(P)\gamma^1 S(P)\\
-\gamma^2 &= S^{-1}(P)\gamma^2 S(P)\\
-\gamma^3 &= S^{-1}(P)\gamma^3 S(P)
\end{aligned} \tag{3.68}$$

we are looking for an $S(P)$ that satisfies all the conditions in Eqn. 3.68. Multiplying Eqns. 3.68 on the left by $S(P)$ and rearranging, its easy to see that $\{S(P), \vec{\gamma}\} = 0$ and $[S(P), \gamma^0] = 0$. The second equation here gives us a hint, γ^0 commutes with itself. A general choice for $S(P)$ is therefore $e^{i\phi}\gamma^0$ since the phase factor doesnt change any physics. Using P to denote $S(P)$ from now on we can write:

$$\begin{aligned}
S(P) &= P = e^{i\phi}\gamma^0\\
\psi_P(x) &= P\psi(x)
\end{aligned} \tag{3.69}$$

Where $\psi_P(x)$ in Eqn. 3.69 above denotes the parity transformed wave function.

$$\begin{aligned}
[P, \gamma^0] = 0 &\rightarrow P\gamma^0 = \gamma^0 P\\
\{P, \vec{\gamma}\} = 0 &\rightarrow -P\vec{\gamma} = \vec{\gamma}P
\end{aligned} \tag{3.70}$$

It is important to note that positive and negative energy solutions derived in 2.31-2.32 are eigenvectors of P with eigenvalues ± 1 respectively. Note that $P^2 = PP = \mathbb{I}$ and so P is its own inverse. At this point we consider the following:

i. Under the parity transformation all spatial parts of vectors will change orientation. Since $\vec{x} \to -\vec{x}$ and $\vec{p} \to -\vec{p}$ the Lorentz invariant scalar product in the exponent of a free wave will not change sign i.e. $p_\mu x^\mu = p_0 x^0 - \vec{p}\cdot\vec{x} \to= p_0 x^0 - (-\vec{p})\cdot(-\vec{x}) = p_0 x^0 - \vec{p}\cdot\vec{x}$.

ii. The gradient operator in the Dirac equation will change sign too as we just mentioned, hence $\nabla \to -\nabla$

iii. The parity transformation will also mean that the spatial part of the electromagnetic 4-current density will be transformed $\vec{J} \to -\vec{J}$. Thus via $\Box A^\mu = 4\pi J^\mu$ the vector potential will also change sign $\vec{A} \to -\vec{A}$. What does this mean for the Dirac equation in the presence of a vector potential ?

We may ask ourselves the question what is the form of the Dirac equation that the parity transformed wave function $\psi_P(x)$ will satisfy ? To answer this question we write down the Dirac equation in the presence of an electromagnetic field

$$(i\hbar \not{\partial} - \frac{e}{c} \not{A}(x) - mc)\psi(x) = 0 \tag{3.71}$$

we now write $\psi(x)$ in terms of the parity transformed wave function $\psi_P(x) = P\psi(x)$ or by using $\psi(x) = P^{-1}\psi(x)$.

$$(i\hbar \not{\partial} - \frac{e}{c} \not{A}(x) - mc)P^{-1}\psi_P(x) = 0 \tag{3.72}$$

multiplying this on the left by P and putting in the $\vec{\gamma}$ and γ^0 into Eqn. 3.72 explicitly we have after minimal simplification:

$$\begin{aligned}(i\hbar P\gamma^0 P^{-1}\frac{\partial}{\partial x^0} + P\vec{\gamma}P^{-1}\cdot\vec{\nabla})\psi_P(x)\\ +(-\frac{e}{c}P\gamma^0 P^{-1}A_0 - P\vec{\gamma}P^{-1}\cdot\vec{A} - mc)\psi_P(x) = 0\end{aligned} \tag{3.73}$$

using Eqn. 3.70 it is trivial to see that P can be moved to the left of each of the γ matrices in Eqn. 3.73 picking up another $-$ sign associated with the spatial parts of A and ∇ and we will get

$$(i\hbar \not{\partial} - \frac{e}{c} \not{A}(x) - mc)\psi(x)_P = 0 \tag{3.74}$$

in other words the parity transformed wave function satisfies the same Dirac equation as the untransformed one.

2 . We now consider the operation of charge conjugation. Since we know the positron exists (!) it becomes obvious that we could very well write our Dirac equation in terms of positrons-and then electrons would be taken as negative energy solutions. For this reason we define a charge conjugated wave function, $\psi_C(x)$ which satisfies the Dirac equation with the sign of the charge e reversed. To find the transformation linking $\psi(x)$ to $\psi_C(x)$ we start by writing the Dirac equation for an electron in the presence of an electromagnetic field:

$$(i\hbar \not{\partial} - \frac{e}{c} \not{A}(x) - mc)\psi(x) = 0 \tag{3.75}$$

and require that there must be a simple rule to write down an equation

$$(i\hbar \not{\partial} + \frac{e}{c} \not{A}(x) - mc)\psi_C(x) = 0 \tag{3.76}$$

for the positron where the wave function $\psi_C(x)$ is related to $\psi(x)$ by the operation of "charge conjugation" which we will now elaborate. As a first step to go to Eqn. 3.76 from Eqn. 3.75 we see that the relative sign of $i\hbar\gamma^\mu\partial_\mu$ and $e\gamma^\mu A_\mu(x)$ must change, after which we will find the rule for changing $\psi(x) \to \psi_C(x)$. We start by taking complex conjugate of Eqn. 3.75 ($A^\mu(x)$ is always real), this reverses the sign :

$$(-i\hbar\gamma^{\mu *}\partial_\mu - \frac{e}{c}\gamma^{\mu *}A_\mu(x) - mc)\psi^*(x) = 0 \tag{3.77}$$

Here explicit indices signifying 4-vector components or the 4-γ matrices have been reintroduced, the reason will become obvious in a few lines. We now assume now that there exists an operator C such that $\psi_C(x) = C\psi^*(x)$ and we insert a $C^{-1}C$ to the right of $\psi(x)$ in Eqn. 3.77 and multiply Eqn. 3.77 on the left by C we obtain:

$$\begin{aligned}&(-i\hbar C\gamma^{\mu *}C^{-1}\partial_\mu - \frac{e}{c}C\gamma^{\mu *}C^{-1}A_\mu(x) - mcCC^{-1})C\psi^*(x)\\ &= (-i\hbar C\gamma^{\mu *}C^{-1}\partial_\mu - \frac{e}{c}C\gamma^{\mu *}C^{-1}A_\mu(x) - mc)\psi_C(x) = 0\end{aligned} \tag{3.78}$$

Comparing Eqns 3.78 and 3.76 we see that if we can find a C such that $\gamma^\mu = -C\gamma^{\mu *}C^{-1}$ then we can define the charge conjugation operation and obtain Eqn. 3.76. Now in our representation the only γ matrix that is complex is γ^2 it is purely imaginary and $(\gamma^{0,1,3})^* = \gamma^{0,1,3}$. It is easy to see that $C = i\gamma^2$ satisfies our requirements. The reader is encouraged to apply the charge conjugation operator $C = i\gamma_2$ to the complex conjugate to the stationary solutions 2.31-2.32 at the end of Chapter 2.

Thus we define the charge conjugation operation:

$$\psi_C(x) = i\gamma^2\psi^*(x) \tag{3.79}$$

Finally, the reader is asked to note that the Charge Conjugation operation is its own inverse and that $C^2 = CC = \mathbb{I}$.

3 . Another discrete symmetry of the Dirac equation is the time reversal symmetry. The reader is surely familiar with this. Any "movie" of a fundamental process such as a particle scattering will be indistinguishable if played backward. Writing the Dirac equation with the space and time parts clearly shown :

$$\begin{aligned}
&(i\hbar \not\partial - \frac{e}{c} \not A(x) - mc)\psi(x) = \\
&(i\hbar)(\gamma^0\partial_0 + \vec{\gamma}\cdot\vec{\nabla} - \frac{e}{c}\gamma^0 A_0(x) + \frac{e}{c}\vec{\gamma}\cdot\vec{A} - mc)\psi(x) = 0
\end{aligned} \tag{3.80}$$

We have to reverse the sign of the time co-ordinate in this expression and rewrite the Dirac equation in terms of the time reversed wave function $\psi_T(x)$. We must first understand that in addition to changing the sign of the time co-ordinate in the Dirac equation we must consider the effect of time reversal on the 4-vector potential $A^\mu(x)$. Recall from Chapter 1 that the four vector potential satisfies:

$$\Box A^\mu(x) = 4\pi J^\mu(x) \tag{3.81}$$

The spatial part of $J^\mu(x)$, $\vec{J}(x)$ represents the flow of charge per unit time per unit area, thus a reversal of the direction of time means that $\vec{J}(x) \to -\vec{J}(x)$ and as a consequence of Eqn. 3.81 $\vec{A}(x) \to -\vec{A}(x)$. Taking everything into account we write:

$$\begin{aligned}
&(-i\hbar\gamma^0\partial_0 + i\hbar\vec{\gamma}\cdot\vec{\nabla})\psi_T(x) \\
&+ (-\frac{e}{c}\gamma^0 A_0(x) + \frac{e}{c}\vec{\gamma}\cdot\vec{A} - mc)\psi_T(x) = 0
\end{aligned} \tag{3.82}$$

We now note also that under time reversal the spatial component of the 4-momentum will also change sign since it is $\propto \frac{dx}{dt}$

and $t \to -t$ implies $\frac{dx}{dt} \to \frac{dx}{d(-t)} = -\frac{dx}{dt}$. Thus the phase of a freely propagating wave under this transformation will change from $\pm i\frac{p_\mu x^\mu}{\hbar} \to \mp i\frac{p_\mu x^\mu}{\hbar}$. Using this hint we complex conjugate Eqn. 3.80 as a first step

$$\begin{aligned}(-i\hbar)(\gamma^{0*}\partial_0 + \vec{\gamma}^* \cdot \vec{\nabla})\psi^*(x) \\ -\frac{e}{c}\gamma^{0*}A_0(x) + \frac{e}{c}\vec{\gamma}^* \cdot \vec{A} - mc)\psi^*(x) = 0\end{aligned} \tag{3.83}$$

Assuming the existence of an operator T such that $T\psi^*(x) = \psi_T(x)$ and hence using $T^{-1}\psi_T(x)$ in place of $\psi^*(x)$ in Eqn. 3.83, multiplying on the left by T, recalling that γ^0 is real and simplifying a little bit we obtain:

$$\begin{aligned}(-i\hbar)(T\gamma^0 T^{-1}\partial_0 + \vec{\gamma}^* \cdot \vec{\nabla})\psi_T(x) \\ -\frac{e}{c}T\gamma^0 T^{-1}A_0(x) + \frac{e}{c}T\vec{\gamma}^* T^{-1} \cdot \vec{A} - mc)\psi_T(x) = 0\end{aligned} \tag{3.84}$$

Note that $\gamma^{2*} = -\gamma^2$ and that all the other gamma matrices are real. Comparing thus Eqn. 3.84 to we obtain the following conditions

$$\begin{aligned}T\gamma^0 T^{-1} &= \gamma^0 \\ T\gamma^{1,3}T^{-1} &= -\gamma^{1,3} \\ T\gamma^2 T^{-1} &= \gamma^2\end{aligned} \tag{3.85}$$

Multiplying on the left with T (and thus removing T^{-1}) and rearranging, we are left with the following requirements that T must satisfy:

$$\begin{aligned}[T, \gamma^0] &= 0 \\ \{T, \gamma^{1,3}\} &= 0 \\ [T, \gamma^2] &= 0\end{aligned} \tag{3.86}$$

Its easy to see that $\gamma^1\gamma^3$ is a valid choice for T.

Finally the reader is asked to consider the application of all three transformations on a wave function

$$\begin{aligned}&T\psi(x) = \gamma^1\gamma^3\psi(x)^* \\ &PT\psi(x) = \gamma^0\gamma^1\gamma^3\psi(x)^* \\ &CPT\psi(x) = i\gamma^2(\gamma^0\gamma^1\gamma^3\psi(x)^*)^* = i\gamma^2\gamma^0\gamma^1\gamma^3\psi(x)\end{aligned} \tag{3.87}$$

Recalling the definition of the γ^5 matrix in equation 3.60 we see that we can simply swap the positions of first γ^0 and γ^2 and then γ^2 and γ^1 incurring a $-$ sign each time to obtain

$$\psi(x)_{CPT} = i\gamma^0\gamma^1\gamma^2\gamma^3\psi(-x) = \gamma^5\psi(x) \tag{3.88}$$

Thus the CPT transformation applied to for example an electron spinor will reverse its charge and its space time trajectory but we see from equation 3.88 that this is simply the electron wave function multiplied by a factor of γ^5. CHECK PLEASE IS THIS ACTUALLY CORRECT ?

3.6 Plane wave solutions of the Dirac Equation by using a Lorentz transformation

In this section we will use our rule for boosting spinors to come up with a plane wave solution of a Dirac particle moving in an arbitrary direction by applying an *appropriate* boost.

Before proceeding we need to revisit Lorentz transformations: until this point we have been using Lorentz transformations to define the linear relationship between the co-ordinates of the same event in a frame x and a frame x' moving along a particular direction with a velocity $+\vec{\beta}$. We want to *use* a Lorentz transformation to come up with a wave function of a moving particle when we are in a stationary frame viewing it as it goes by. We do not want a Lorentz transformation that represents a view of the particle from another frame that moves with velocity $+\vec{\beta}$ with respect to us, therefore we will create a boost matrix S representing a boost of $-\vec{\beta}$ and apply it to ψ, so then we will view the particle as it flies past us with velocity $+\vec{\beta}$-what we expect from the convention of a plane wave solution. We will make the change $\vec{\beta} \to -\vec{\beta}$ at the very end of our calculation.

We will first calculate a boost for a particle that moves in an arbitrary direction with 4-momentum $(\frac{E}{c}, \vec{p})$ and the direction of its propagation is given by the three direction cosines $\hat{n} = (\cos\alpha_1, \cos\alpha_2, \cos\alpha_3)$ and of course $\hat{n}\cdot\hat{n} = \cos\alpha_1^2 + \cos\alpha_2^2 + \cos\alpha_3^2 = 1$.

Recall from chapter 1 section 1.8 that we can make an infinitesimal boost by using the $I^\mu{}_\nu$ matrices and infinitesimal velocities. The equations 1.45-1.47 define the I arrays for the x, y and z directions. For a plane wave propagating along $\hat{n}$ we have $\Delta\omega$ as the absolute value of an infinitesimal velocity, the components of velocities along x, y and z are then given simply by $\Delta\omega\cos\alpha_1, \Delta\omega\cos\alpha_2$ and $\Delta\omega\cos\alpha_3$. The matrix $\epsilon^{\mu}_{\nu}(\hat{n})$ is then $\Delta\omega(\cos\alpha_1 I^{\mu}{}_{\nu}(x) + \cos\alpha_2 I^{\mu}{}_{\nu}(x) + \cos\alpha_3 I^{\mu}{}_{\nu}(z))$. We take the term in the brackets to define $I^{\mu}_{\nu}(\hat{n})$

$$I^{\mu}{}_{\nu}(\hat{n}) = \cos\alpha_1 I^{\mu}{}_{\nu}(x) + \cos\alpha_2 I^{\mu}{}_{\nu}(y) + \cos\alpha_3 I^{\mu}{}_{\nu}(z) \tag{3.89}$$

represented as a matrix array Eqn. 3.89 becomes

$$I^{\mu}{}_{\nu}(\hat{n}) = \begin{pmatrix} 0 & -\cos\alpha_1 & -\cos\alpha_2 & -\cos\alpha_3 \\ -cos\alpha_1 & 0 & 0 & 0 \\ -\cos\alpha_2 & 0 & 0 & 0 \\ -\cos\alpha_3 & 0 & 0 & 0 \end{pmatrix} \tag{3.90}$$

Note that $I^{\mu\nu}$ (one index raised) has non zero entries $I^{01} = -I^{10} = +\cos\alpha_1$, $I^{02} = -I^{20} = +\cos\alpha_2$ and $I^{03} = -I^{30} = +\cos\alpha_3$, (the reader is asked to review the material leading to Eqns 1.49-1.50).

We now recall that Eqn. 3.25 defines an infinitesimal boost

$$S = \mathbb{I} - \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta} = \mathbb{I} + \frac{1}{8}\Delta\omega[\gamma_0, \gamma_j]I^{0j}(\hat{n}) + \frac{1}{8}\Delta\omega[\gamma_j, \gamma_0]I^{j0}(\hat{n}) \tag{3.91}$$

where $j = 1, 2, 3$.

$$S = \mathbb{I} - \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta} = \mathbb{I} + \frac{1}{8}(\Delta\omega)([\gamma_0, \gamma_j]I^{0j}(\hat{n}) + [\gamma_j, \gamma_0]I^{j0}(\hat{n})) \tag{3.92}$$

Using the elements of $I^{\mu\nu}(\hat{n})$ and $I^{\mu\nu}(\hat{n}) = -I^{\nu\mu}(\hat{n})$ and $\gamma_j\gamma_0 = -\gamma_0\gamma_j$ we obtain

$$\sigma_{\mu\nu}I^{\mu\nu} = 2[\gamma_0, \gamma_j]I^{0j}(\hat{n}) = 2\cos\alpha_j(\gamma_0\gamma_j - \gamma_j\gamma_0) = 4\cos\alpha_j\gamma_0\gamma_j \tag{3.93}$$

where the repeated index j implies summation. Substituting in Eqn. 1.52 we obtain:

$$\lim_{N\to\infty}(\mathbb{I} + \frac{\omega}{2(N)}\cos\alpha_j\gamma_0\gamma_j)^N \tag{3.94}$$

Recall now that these are the γ_j with the indices *lowered* thus are thus $\gamma_0\gamma_j = -\gamma^0\gamma^j$. The product in equation 3.94 is by the definition of the exponential :

$$e^{\frac{-\omega}{2}\cos\alpha_j\gamma^0\gamma^j} \tag{3.95}$$

The series $e^{\frac{-\omega}{2}\cos\alpha_j\gamma^0\gamma^j}$ can be expanded out:

$$e^{\frac{-\omega}{2}\cos\alpha_j\gamma^0\gamma^j} = \mathbb{I} - \frac{\omega}{2}\cos\alpha_j\gamma^0\gamma^j + \frac{\omega^2}{4\times 2!}(\cos\alpha_j\gamma^0\gamma^j)^2 - \frac{\omega^3}{8\times 3!}(\cos\alpha_j\gamma^0\gamma^j)^3 + \cdots \tag{3.96}$$

This is actually quite easy to simplify, looking at the $(\cos\alpha_j\gamma^0\gamma^j)^2$ term this is simply $(\cos\alpha_j\gamma^0\gamma^j)(\cos\alpha_k\gamma^0\gamma^k) = (\cos\alpha_j)(\cos\alpha_j)\mathbb{I} + \cos\alpha_j\cos\alpha_k\{\gamma^j,\gamma^k\}$ here the anticommutator contains all terms with $j \neq k$ and so equals zero, $\cos\alpha_j\cos\alpha_j$ is the *sum* $\cos\alpha_1^2 + \cos\alpha_2^2 + \cos\alpha_3^2 = 1$. That means in any term we can group together all the factors $(\cos\alpha_j\gamma^0\gamma^j)^2$ and write $\mathbb{I}$–thus simplifying the terms containing even powers of $\cos\alpha_j\gamma^0\gamma^j)$, for those containing odd powers of $\cos\alpha_j\gamma^0\gamma^j$ we will of course obtain $\cos\alpha_j\gamma^0\gamma^j$, remembering the powers of $\frac{\omega}{2}$ we obtain:

$$e^{\frac{-\omega}{2}\cos\alpha_j\gamma^0\gamma^j} = \mathbb{I} - \frac{\omega}{2}\cos\alpha_j\gamma^0\gamma^j + \frac{\omega^2}{4\times 2!}\mathbb{I} - \frac{\omega^3}{8\times 3!}\cos\alpha_j\gamma^0\gamma^j + \cdots \tag{3.97}$$

the even powers of the $\frac{\omega}{2}$ terms are identified as the Taylor expansion of $\cosh\frac{\omega}{2}$ and the odd with $\sinh\frac{\omega}{2}$ the $\gamma^0\gamma^k$ terms associated with the $\sinh\frac{\omega}{2}$. The product $\gamma^0\gamma^j$ is equal to $\left(\begin{array}{c|c} 0 & \sigma_j \\ \hline \sigma_j & 0 \end{array}\right)$

We have therefore:

$$\cosh\frac{\omega}{2}\left(\begin{array}{c|c} \mathbb{I} & 0 \\ \hline 0 & \mathbb{I} \end{array}\right) - \sinh\frac{\omega}{2}(\cos\alpha_1\left(\begin{array}{c|c} 0 & \sigma_1 \\ \hline \sigma_1 & 0 \end{array}\right) + \cos\alpha_2\left(\begin{array}{c|c} 0 & \sigma_2 \\ \hline \sigma_2 & 0 \end{array}\right) + \cos\alpha_3\left(\begin{array}{c|c} 0 & \sigma_3 \\ \hline \sigma_3 & 0 \end{array}\right)) \tag{3.98}$$

Before simplifying this into an explicit form we write the cosh and sinh in terms of the momentum and energy. Note that in the very first chapter we had defined $\sinh\omega = \gamma\beta$ and $\cosh\omega = \gamma$ for a parameterization of a Lorentz boost. We have to be a little careful here since we have several components of velocity. It should be clear that the boost represented by $\tanh\omega = \beta$ has the same meaning in all frames and doesn't depend on orientation of the velocity, from this we trivially obtain $\cosh\omega = \frac{1}{\sqrt{1-\tanh^2\omega}} = \gamma$.

We then use the following well known relations:

$$\cosh^2\frac{\omega}{2} - \sinh^2\frac{\omega}{2} = 1 \tag{3.99}$$

and

$$\cosh\omega = \cosh^2\frac{\omega}{2} + \sinh^2\frac{\omega}{2} \tag{3.100}$$

to write down

$$1 + \cosh\omega = 2\cosh^2\frac{\omega}{2} \tag{3.101}$$

and finally using $\cosh\omega = \gamma$ from above we can write

$$1 + \gamma = 2\cosh^2\frac{\omega}{2} \tag{3.102}$$

and using $\gamma = \frac{E}{mc^2}$ have :

$$\cosh\left(\frac{\omega}{2}\right) = \sqrt{\frac{E + mc^2}{2mc^2}} \tag{3.103}$$

Now to tackle the $\sinh\frac{\omega}{2}$, we use Eqn. 3.99 and Eqn. 3.103 to write:

$$\sinh^2\frac{\omega}{2} = \frac{E + mc^2}{2mc^2} - 1$$

This is simply:

$$\begin{aligned}
\sinh^2\frac{\omega}{2} &= \frac{E + mc^2}{2mc^2} - \frac{E + mc^2}{E + mc^2} \\
&= \frac{E^2 + 2mc^2E + m^2c^4 - 2mc^2E - 2mc^2c^4}{2mc^2(E + mc^2)} \\
&= \frac{E^2 - m^2c^4}{2mc^2(E + mc^2)} = \frac{\vec{p}\cdot\vec{p}c^2}{2mc^2(E + mc^2)}
\end{aligned}$$

Which finally brings us to:

$$\sinh\left(\frac{\omega}{2}\right) = \frac{\mid\vec{p}\mid c}{\sqrt{2mc^2(E + mc^2)}} \tag{3.104}$$

We are now ready to use the fact that we want to apply the boost in the opposite direction for which we simply reverse the signs of the direction cosines in the expression for the boost matrix in Eqn. 3.98 followed by substituting for $\cosh\frac{\omega}{2}$ and $\sinh\frac{\omega}{2}$. Using $p_{x,y,z} = \mid\vec{p}\mid\cos\alpha_{1,2,3}$ we obtain :

$$\begin{aligned}
&\sqrt{\frac{E + mc^2}{2mc^2}}\Big(\left(\begin{array}{c|c} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{array}\right) + \frac{cp_x}{E + mc^2}\left(\begin{array}{c|c} 0 & \sigma_1 \\ \sigma_1 & 0 \end{array}\right) + \frac{cp_y}{E + mc^2}\left(\begin{array}{c|c} 0 & \sigma_2 \\ \sigma_2 & 0 \end{array}\right) + \\
&\qquad \frac{cp_z}{E + mc^2}\left(\begin{array}{c|c} 0 & \sigma_3 \\ \sigma_3 & 0 \end{array}\right)\Big)
\end{aligned} \tag{3.105}$$

combining this into one big matrix and using the Pauli matrices we get the full boost matrix for an arbitrary boost:

$$\sqrt{\frac{E+mc^2}{2mc^2}}\begin{pmatrix} 1 & 0 & \frac{cp_z}{E+mc^2} & \frac{cp_-}{E+mc^2} \\ 0 & 1 & \frac{cp_+}{E+mc^2} & \frac{-cp_z}{E+mc^2} \\ \frac{cp_z}{E+mc^2} & \frac{cp_-}{E+mc^2} & 1 & 0 \\ \frac{cp_+}{E+mc^2} & \frac{-cp_z}{E+mc^2} & 0 & 1 \end{pmatrix} \tag{3.106}$$

where $p_\pm = p_x \pm ip_y$.

Now the boost matrix 3.106 must be applied to the stationary wave functions in Eqns. 2.31-2.32 however we have not yet discussed the exponential factors for the plane waves: for the stationary solutions in Chapter 2 these were $e^{\pm i\frac{mc^2t}{\hbar}}$. The correct factors for a moving particle require just a bit of thought. We note that the exponential factor is a Lorentz invariant (see discussion preceding 1.32), if we denote the plane wave factors $e^{-\epsilon_r\frac{imc^2t}{\hbar}}$ with $\epsilon_r = 1$ for $r = 1{,}2$ and $\epsilon_r = -1$ for $r = 3{,}4$ thus representing the positive and negative energy solutions first encountered in Chapter 2 in Eqns. 2.31-2.32 then it is clear that when viewed in any frame of reference we need to replace the exponential plane wave factors with $e^{-\epsilon_r i\frac{p_\mu x^\mu}{\hbar}}$. We will write these out explicitly and remind the reader that $p_\mu = (p_0, -p_x, -p_y, -p_z) = (\frac{E}{c}, -p_x, -p_y, -p_z)$ and $x^\mu = (x_0, x^1, x^2, x^3) = (ct, x, y, z)$. Defining $\psi_{1,2} = u_{1,2}(p)e^{-i\frac{p_\mu x^\mu}{\hbar}}$ and $\psi_{3,4} = v_{1,2}(p)e^{i\frac{p_\mu x^\mu}{\hbar}}$ we have the following plane wave solutions as a function of x for the positive energy solutions where $N = \sqrt{\frac{E+mc^2}{2mc^2}}$:

$$\psi_1(x) = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \end{pmatrix} e^{-i\frac{Et-\vec{p}\cdot\vec{x}}{\hbar}}, \quad \psi_2(x) = N\begin{pmatrix} 0 \\ 1 \\ \frac{cp_-}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \end{pmatrix} e^{-i\frac{Et-\vec{p}\cdot\vec{x}}{\hbar}}, \tag{3.107}$$

and the following for the negative energy solutions:

$$\psi_3(x) = N\begin{pmatrix} \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \\ 1 \\ 0 \end{pmatrix} e^{i\frac{Et-\vec{p}\cdot\vec{x}}{\hbar}}, \quad \psi_4(x) = N\begin{pmatrix} \frac{cp_-}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \\ 0 \\ 1 \end{pmatrix} e^{i\frac{Et-\vec{p}\cdot\vec{x}}{\hbar}} \tag{3.108}$$

These can be summarized in the fairly standard notation [7] [8] as $\omega_r(p)e^{-\epsilon_r\frac{ip\cdot x}{\hbar}}$ with $\epsilon_r = 1$ for $r = 1, 2$ and -1 for $r = 3, 4$ and we have:

$$\omega_1(p) = N \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \end{pmatrix}, \quad \omega_2(p) = N \begin{pmatrix} 0 \\ 1 \\ \frac{cp_-}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \end{pmatrix}, \tag{3.109}$$

$$\omega_3(p) = N \begin{pmatrix} \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad \omega_4 b(p) = N \begin{pmatrix} \frac{cp_-}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \\ 0 \\ 1 \end{pmatrix} \tag{3.110}$$

3.7 Spinor rotation

Let us now consider rotations of the wave function $\psi(x)$. We begin by noting that rotations leave the Lorentz invariant quantity $c^2 - \vec{x}\cdot\vec{x}$ unchanged since the norm of the spatial remains unchanged. The reader is asked to recall the derivation of the Lorentz boost for a spinor. In that spirit we first consider a matrix representing an infinitesimal rotation about the z axis:

$$\epsilon^{\nu}_{\mu,R}(z) = \Delta\phi_z \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \Delta\phi_z I^{\nu}_{\mu,R}(z) \tag{3.111}$$

where the subscript $_R$ denotes rotation, the argument z the axis about which this is to be performed and $\Delta\phi_z$ the infinitesimal angle of rotation. It should be easy to see that one can apply an infinite number of such infinitesimal rotations analogous to what was done in Section **??** we will obtain the usual rotation matrix for a rotation about z:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi_z & \sin\phi_z & 0 \\ 0 & -\sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.112}$$

confirming that indeed 3.111 represents an infinitesimal rotation. Note that the time and z axis are unchanged. To apply this to a spinor we note that the only non-zero elements of $I^{\nu\mu}_R(z)$ are $I^{12} = 1$ and $I^{21} = -1$ and using our expression for S and the definition of $\sigma_{\mu\nu}$ we can write down

$$S = \mathbb{I} - \tfrac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta} = \mathbb{I} + \tfrac{1}{8}(\Delta\phi_z)([\gamma_1,\gamma_2] - [\gamma_2,\gamma_1]) \tag{3.113}$$

$$= \mathbb{I} + \tfrac{1}{4}\Delta\phi_z[\gamma_1\gamma_2] = \mathbb{I} + \tfrac{1}{2}\Delta\phi_z\gamma_1\gamma_2 \tag{3.114}$$

It is easily verified that

$$\gamma_1\gamma_2 = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} = -i\left(\begin{array}{c|c} \sigma_3 & 0 \\ \hline 0 & \sigma_3 \end{array}\right) \tag{3.115}$$

in fact in general for spatial indices i, j it's easily verified that

$$\gamma^i\gamma^j = -i\epsilon_{ijk}\Sigma^k \tag{3.116}$$

We denote the resulting matrix on the right hand side of Eqn. 3.115 by Σ_3 and can once again derive that infinite applications (as in Section 3.6)of this will result in a finite rotation represented by the 4×4 matrix:

$$e^{i\frac{\phi_z}{2}\Sigma_3} \tag{3.117}$$

which we can then generalize to a rotation about an arbitrary axis represented by $\hat{n}$

$$e^{i\frac{\phi}{2}\hat{n}\cdot\vec{\Sigma}} \tag{3.118}$$

recognizing a very familiar expression for the rotation of spinors we note that the $\phi \to \phi+2\pi$ yields an overall factor of -1 due to the factor of $\frac{1}{2}$ and the resulting factor of $i\pi$ in the exponent. Only a rotation of $\phi \to \phi+4\pi$ will restore the spinor to itself. We can see that this rotation can be compared directly to that for a Pauli spinor. Finally we point out that we have not replaced ϕ by $-\phi$ as we did for $\vec{\beta}$ in Section 3.6 since we are only trying to make a comment on a particular transformation property and not derive a wave function. The 4 degrees of freedom therefore are confirmed to represent spin for the positive and negative energy solutions. We had assumed this in previous chapters but have now confirmed this due to 4π rotation property.

3.8 Normalization, spin and energy projection operators

We will examine some of the properties of the spinors listed in Eqn. 3.107-3.108. We begin with the normalization of the spinors, recall that the Hermitian conjugate of the spinors is written as a row matrix containing the complex conjugate of the entries. As an example recall $\omega_1(p)$:

$$\omega_1(p) = \sqrt{\frac{E+mc^2}{2mc^2}} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \end{pmatrix} \tag{3.119}$$

the Hermitian conjugate of this is:

$$\omega_1^\dagger(p) = \sqrt{\frac{E+mc^2}{2mc^2}} \begin{pmatrix} 1 & 0 & \frac{cp_z}{E+mc^2} & \frac{cp_-}{E+mc^2} \end{pmatrix} \tag{3.120}$$

where $(p_+)^* = (p_x - ip_y) = p_-$ has been used. It is easily shown that $\omega_1^\dagger(x)u_1(x) = \frac{E^2+m^2c^4+2Emc^2+c^2p^2}{2mc^2(E+mc^2)} = \frac{E}{mc^2}$ and the reader can verify the following orthogonality relations for all of the ω_i

$$\omega_r^\dagger(p)\omega_{r'}(p) = \delta_{rr'}\epsilon_r \frac{E}{mc^2} \tag{3.121}$$

where $\epsilon_r = +1$ for $r = 1, 2$ ($E > 0$ solutions) and $\epsilon_r = -1$ for $i = 3, 4$ ($E < 0$ solutions).

The reader is asked to note that $\frac{E}{mc^2} = \gamma$, the quantity $\psi^\dagger\psi$ is a probability density and is not a Lorentz invariant quantity, it is in fact the zeroth component of the probability four current $\psi^\dagger\gamma^0\gamma^\mu\psi$ as defined and verified in Eqn. 3.13. The integral of $\psi^\dagger\psi$ over all space is the probability of finding the particle represented somewhere and should equal to 1, the factor of γ will be exactly cancelled by the factor of $\frac{1}{\gamma}$ that the volume element of the integral acquires when observed in an arbitrary frame of reference.

Recall that in the adjoint spinor was first defined in Eqn. 3.14 and made several appearances in section 3.4 we had defined the adjoint spinor $\bar{\psi} = \psi^\dagger\gamma^0$, we do the same here, thus as a specific example :

$$\bar{\omega_1}(p) = \omega_1^\dagger(p)\gamma^0 = \sqrt{\frac{E+mc^2}{2mc^2}} \begin{pmatrix} 1 & 0 & \frac{cp_z}{E+mc^2} & \frac{cp_-}{E+mc^2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.122}$$

which is simply:

$$\bar{\omega_1}(p) = u_1^\dagger(p)\gamma^0 = \sqrt{\frac{E+mc^2}{2mc^2}}\left(1\;0\;-\tfrac{cp_z}{E+mc^2}\;-\tfrac{cp_-}{E+mc^2}\right) \tag{3.123}$$

using this specific example one can easily verify that (for example) $\bar{\omega_1}\omega_1 = 1$ and also that :

$$\bar{\omega_r}(p)\omega_{r'}(p) = \delta_{rr'}\epsilon_r \tag{3.124}$$

written in terms of ψ these are

$$\bar{\psi}_r(x)\psi_{r'}(x) = \epsilon_r\delta_{rr'} \tag{3.125}$$

Here in Eqn. 3.125 the orthoganality relations of the spinors and their adjoints are summarized.

Note that we can now use the Dirac equation to obtain the Dirac equation in momentum space, which doesnt include the exponential plane wave factor. We write down the Dirac equation separately for the $E > 0$ and $E < 0$ wave functions:

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi_{r=1,2}(x) = 0 \tag{3.126}$$

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi_{r=3,4}(x) = 0 \tag{3.127}$$

by operating the derivative operators on the exponential term we obtain after canceling the exponential factor:

$$\begin{aligned}(\not{p} - mc)\omega_{1,2}(p) = 0\\ (\not{p} + mc)\omega_{3,4}(p) = 0\end{aligned} \tag{3.128}$$

each of these equations is a matrix equation satisfied by the spinors in equations 3.107-3.108, note that $\not{p} = \gamma^0p^0 - \vec{\gamma}\cdot\vec{p} = \gamma^0 - \gamma^1p^1 - \gamma^2p^2 - \gamma^3p^3$.

We now recall that the free wave solutions to the Dirac equation in equations 3.109 3.110 were obtained from appropriately boosting the stationary state solutions in equations 2.30 2.31. At that point we had noted that these were spin eigenfunctions, this is however clearly not true for plane wave solutions propagating in an arbitrary direction as one can easily verify. This is easy to understand, now that they have been boosted, the unit vector $\hat{s} = (s_x, s_y, s_z) = (0, 0, 1)$ pointing in the direction of spin must now be considered in a different Lorentz frame. It is better therefore to define the unit 4-vector $s^\mu = (s_0, s_x, s_y, s_z)$, it is clear that $s^\mu s_\mu = -1$ is a Lorentz invariant, we know this by taking the scalar product of s^μ with itself in the rest frame where it has a single purely spatial component. It is also easy to see that since in the rest frame $p^\mu = (\frac{E}{c}, 0, 0, 0)$, the scalar product $p^\mu s_\mu = 0$ is Lorentz invariant. We refer to s^μ as a the spin polarization vector. As mentioned a few lines back, the freely propagating wave functions are not eigensfunctions of spin and so it is more meaningful to speak of the projections of their spin along a particular direction of polarization. So we'll now define the spin-projection operators which do just that. We write down (rather arbitrarily)

$$\frac{1 \pm \gamma^5 \not{s}}{2} \tag{3.129}$$

claiming that this is an appropriate representation of the spin projection operator, where $\not{s} = \gamma^\mu s_\mu$. Now it is very to verify that

$$\gamma^5 \gamma^k = \begin{pmatrix} \sigma_k & 0 \\ 0 & -\sigma_k \end{pmatrix}$$

we now go to the rest frame to examine the action of the operator that we have just constructed, thus setting $s_\mu = (0, 0, 0, -1)$ we evaluate the spin-up projection operator $\frac{1 \pm \gamma^5 \not{s}}{2}$ this is easy to write down:

$$\begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}$$

It is very easy to see that this projects out the spin up stationary, $E > 0$ state in equation 2.30 and its action on the spin down stationary $E > 0$ state yields 0. However we see that this projection operator annihilates the *what we thought was the spin up* $E < 0$ *stationary state* in 2.31 and projects out *what we thought was the spin down* $E < 0$ *stationary solution.* In fact this is fine ! In chapter 2, at the end of section 2.4 in a discussion of how an unoccupied $E < 0$ state would appear to us we had concluded that it would have spin, charge and momentum opposite in sign to the state itself, so the operator defined in equation 3.129 is a valid description of what we expect to observe. This in fact will be how we describe positrons. We now denote the positive energy spinors $\omega_1(p), \omega_2(p)$ by $u(p, s)$ or $u_s(p)$ with $s = 1, 2$ respectively and $\omega_4(p), \omega_3(p)$ by $v(p, s)$ or $v_s(p)$ with $s = 1, 2$ (note the order!) in keeping with our intepretation of negative energy states and the orientation of their spin in the rest frame.

All the previous equations involving $\omega_r(p)$ can be rewritten in terms of $u_{1,2}(p) = \omega_{1,2}(p)$ and $v_{1,2}(p) = \omega_{4,3}(p)$. Having redefined our solutions in the light of examing the action of the spin projection operator, we can (very easily) rewrite the relations in equation 3.130 as we have below

$$\begin{aligned}(\not{p} - mc)u_{1,2}(p) &= 0 \\ (\not{p} + mc)v_{1,2}(p) &= 0\end{aligned} \tag{3.130}$$

Using Eqn. 3.130 we can easily see that

$$\begin{aligned}\not{p}u_{1,2}(p) &= mcu_{1,2} \\ \text{and also that } -\not{p}v_{1,2}(p) &= mcv_{1,2}\end{aligned} \tag{3.131}$$

from these two it is easy to see that

$$(\frac{\not{p} + mc}{2mc})u_{1,2}(p) = \frac{(mc + mc)u_{1,2}(p)}{2mc} = u_{1,2}(p) \tag{3.132}$$

and

$$(\frac{-\not{p} + mc}{2mc})v_{1,2}(p) = v_{1,2}(p) \tag{3.133}$$

we also note that since

$$\frac{\not{p} + mc}{2mc} + \frac{-\not{p} + mc}{2mc} = \mathbb{I} \tag{3.134}$$

thus $\frac{\not{p}+mc}{2mc}$ and $\frac{-\not{p}+mc}{2mc}$ project out positive and negative energy states and sum to identity. So we define the projection operators for positive and negative energy states of the Dirac equation

$$\begin{aligned}\Lambda_+ &= \frac{\not{p} + mc}{2mc} \\ \Lambda_- &= \frac{-\not{p} + mc}{2mc}\end{aligned} \tag{3.135}$$

the only thing that has not been validated is that $\Lambda_+ v_{1,2} = 0$ and $\Lambda_- u_{1,2} = 0$ which is left to the reader to verify. This would complete the justification for the $\Lambda_\pm$ being projection opertators. The reader is also asked to check using the orthogonality relations 3.121 that :

$$\begin{aligned}\Lambda_+\gamma^0 u_{1,2}(p) &= \frac{E}{mc^2}u_{1,2}(p) \\ \Lambda_-\gamma^0 u_{1,2}(p) &= 0 \\ \Lambda_-\gamma^0 v_{1,2}(p) &= \frac{E}{mc^2}v_{1,2}(p) \\ \Lambda_+\gamma^0 v_{1,2}(p) &= 0\end{aligned} \tag{3.136}$$

We will now define operators that will be extremely useful in averaging or summing over all possible spin states in scattering processes where particle spins are not observed. These sum over spin states of the spinors is defined for the $E > 0$ and $E < 0$ solutions. We will do this by forming the *matrices*:

$$\sum_{s=1}^{2} u_s(p)\bar{u}_s(p) = u_1(p)\bar{u}_1(p) + u_2(p)\bar{u}_2(p)$$

expanded explicitly these are:

$$N^2(\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \end{pmatrix} \left(1\; 0\; -\frac{cp_z}{E+mc^2}\; -\frac{cp_-}{E+mc^2}\right) + \begin{pmatrix} 0 \\ 1 \\ \frac{cp_-}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \end{pmatrix} \left(0\; 1\; -\frac{cp_+}{E+mc^2}\; \frac{cp_z}{E+mc^2}\right))$$

where $N = \sqrt{\frac{E+mc^2}{2mc^2}}$. Multiplying the adjoint spinors (rows) by each entry in the spinors to form each row in the resulting matrix we obtain:

$$\begin{pmatrix} \frac{E+mc^2}{2mc^2} & 0 & \frac{-p_z}{2mc} & \frac{-p_-}{2mc} \\ 0 & \frac{E+mc^2}{2mc^2} & \frac{-p_+}{2mc} & \frac{p_z}{2mc} \\ \frac{p_z}{2mc} & \frac{p_-}{2mc} & \frac{-c^2(p_z^2+p_+p_-)}{2mc^2(E+mc^2)} & 0 \\ \frac{p_+}{2mc} & \frac{-p_z}{2mc} & 0 & \frac{-c^2(p_z^2+p_+p_-)}{2mc^2(E+mc^2)} \end{pmatrix} \tag{3.137}$$

the above expression can easily be simplified further using:

$$-(c^2p_z^2 + p_-p_+) = -c^2 \mid \vec{p}^2 \mid = -(E^2 - mc^4) = -(E + mc^2)(E - mc^2)$$

and

$$E \pm mc^2 = c(\frac{E}{c} \pm mc)$$

The matrix 3.137 can simply be rewritten and easily expressed in terms of

$$\not{p} = \gamma^\mu p_\mu$$

where

$$p_\mu = (p_0, p_1, p_2, p_3) = (\frac{E}{c}, -p_x, -p_y, -p_z)$$

using this we get:

$$\begin{pmatrix} \frac{E+mc^2}{2mc^2} & 0 & \frac{-p_z}{2mc} & \frac{-p_-}{2mc} \\ 0 & \frac{E+mc^2}{2mc^2} & \frac{-p_+}{2mc} & \frac{p_z}{2mc} \\ \frac{p_z}{2mc} & \frac{p_-}{2mc} & \frac{-E+mc^2}{2mc} & 0 \\ \frac{p_+}{2mc} & \frac{-p_z}{2mc} & 0 & \frac{-E+mc^2}{2mc} \end{pmatrix} = \frac{\not{p} + mc}{2mc} \tag{3.138}$$

the corresponding spin sum for the $E < 0$ solutions is straightforward to evaluate. Both spin sum relations are summarized below and written in terms of the energy projection operators:

$$\Lambda_+ = \sum_{s=1}^{2} u_{s,\alpha}(p)\bar{u}_{s,\beta}(p) = (\frac{\not{p} + mc}{2mc})_{\alpha\beta} \tag{3.139}$$

$$-\Lambda_- = \sum_{s=1}^{2} v_{s,\alpha}(p)\bar{v}_{s,\beta}(p) = (\frac{\not{p} - mc}{2mc})_{\alpha\beta} = -(\frac{-\not{p} + mc}{2mc})_{\alpha\beta} \tag{3.140}$$

Finally the reader is asked to hearken back to the CPT operation defined in equations 3.87 3.88. With the explicit representation of the γ^5 matrix in equation 3.60 it is trivial to use the plane wave solutions and show that for example :

$$CPT\psi_3(x) = CPT(v_1(p)e^{i\frac{p\cdot x}{\hbar}}) = iu_1(p)e^{i\frac{ip\cdot x}{\hbar}} \tag{3.141}$$

Note that since $\psi_1(x)$ is $u_1(p)e^{\frac{-ip\cdot x}{\hbar}}$, equation 3.141 is showing the result of the CPT operation. Note that the result is an irrelevant phase factor times a positive energy solution but *with its sense of time backward*. The wave function $u_1(p)e^{i\frac{ip\cdot x}{\hbar}}$ has an charge opposite to that of the negative energy electron represented by $\psi(x)_3$ and its sense of time is also backward: this leads us to the conclusion that negative energy electrons running backward in time describe positive energy positrons running forward in time.

3.9 Helicity and Chirality

The reader is asked to recall the solutions of the Dirac equation in momentum space, we rewrite the $E < 0$ and $E > 0$ equations as *eigenvalue* equations:

$$\begin{aligned} i\hbar\,\not{p}u_{1,2}(p) &= mcu^{1,2}(p) \\ i\hbar\,\not{p}v_{1,2}(p) &= -mcv^{1,2}(p) \end{aligned} \tag{3.142}$$

note that for each of these equations, for each E and $\vec{p}$ there are two independent solutions. This of course tells us the the $u(p)$ and $v(p)$ are eigenstates of $\not{p}$ and some other yet to be determined operator. We've already excluded the spin or the angular momentum as one of these operators, as we had calculated the commutator of these with the Dirac Hamiltonian with the spin and orbital angular momenta with a Dirac particle and found that these didn't commute and hence are not conserved.

So let's remind ourselves of the Hamiltonian form of the Dirac equation and write

$$\hat{H}_{DIRAC} = (c\vec{\alpha} \cdot \hat{p}) + \beta mc^2)) \tag{3.143}$$

we now define the operator

$$\hat{h} = \frac{\vec{\Sigma} \cdot \hat{p}}{|\,\vec{p}\,|} \tag{3.144}$$

where $\hat{p}$ is the momentum operator in equation 3.144 and $\vec{\Sigma}$ are the 4-dimensional Pauli matrices first defined in Chapter 1. We'll now calculate the commutator $[\hat{H}_{DIRAC}, \hat{h}]$ and this of course has two pieces with the piece $[\beta mc^2, \hat{h}]$ easily shown to be zero. We calculate the first piece $[c\vec{\alpha}\cdot\hat{p}, \hat{h}]$ using the j^{th} and k^{th} components, its clear that this piece is proportional to:

$$\begin{aligned} [\alpha_j, \Sigma_k]\hat{p}_j\hat{p}_k = \left(\begin{array}{c|c} 0 & [\sigma_j, \sigma_k] \\ \hline [\sigma_j, \sigma_k] & 0 \end{array}\right)\hat{p}_j\hat{p}_k =& 2i\epsilon_{ijk}\left(\begin{array}{c|c} 0 & \sigma_i \\ \hline \sigma_i & 0 \end{array}\right)\hat{p}_j\hat{p}_k \\ =& 2i\epsilon_{ijk}\alpha_i\hat{p}_j\hat{p}_k = 0 \end{aligned} \tag{3.145}$$

thus operator $\hat{h}$ defined in equation 3.144 is seen to commute with the Dirac Hamiltonian. It is clear that this operator represents the scalar product of the spin with the spatial momentum. There is a shortcut we can take to check its eigenvalues, in the plane wave solutions of the Dirac equation 3.109 3.110 we simply set $p_{\pm} = 0$ and consider motion along z only. Its then trivial to see that the solutions to the Dirac equation are eigenfunctions of $\hat{h}$ with eigenvalues ± 1, these are eigenvalues of the *helicity* operator, the name evocative of "helix" is suitable since we are considering the component of spin along the direction of motion of the particle.

Note that the helicity of a massive Dirac particle is not a Lorentz invariant quantity, one need only consider a frame in which the momentum of the particle points opposite to the direction in another frame. Of course this is only possible if the particle has finite, non-zero mass. If on the other hand the particle is massless and travels at the speed of light, there is no frame in which the helicity appears oppositely oriented to its direction in another. The helicity of a particle travelling at the speed of light is referred to as chirality. We can find the chirality operator by noting first that

$$\hat{h} = \frac{\vec{\Sigma} \cdot \hat{p}}{\mid \vec{p} \mid} = \frac{\gamma^5 \gamma^k \gamma^0 \hat{p}_k}{\mid \vec{p} \mid} \tag{3.146}$$

simply by multiplying the matrices, now switching from $\hat{p_k}$ to p_k the *covariant* components of the momentum and using the Dirac equation in momentum space 3.142 we get $\gamma^0 p_0 = \vec{\gamma} \cdot \vec{p}$. A particle travelling at light speed will have $E^2 = c^2 \vec{p} \cdot \vec{p}$ and so $p_0 = \mid \vec{p} \mid$, and so equation 3.146 becomes after substitutng $\gamma^0 p_0$ for $\gamma_k \hat{p}_k$

$$\frac{\gamma^5 \gamma^k \gamma^0 p_k}{\mid \vec{p} \mid} = \frac{\gamma^5 \gamma^0 \gamma^0 p_0}{p_0} = \gamma^5 \tag{3.147}$$

thus γ^5 is identified as the chirality operator, which does not commute with the Dirac Hamiltonian unless the mass of the particle is zero-this is again easily verified using the structure of γ_5 and the α and β matrices.

3.10 Negative energy solutions, Klein Paradox and Zitterbewegung

We now revisit some interesting paradoxes and effects associated with negative energy solutions. One of which manifests itself in considering the familiar problem of a particle impingent on a step potential-but for the case of the Dirac equation. ZITTERBEWEGUNG

We first refresh the reader's memory with the problem of a wave described by the Schrödinger equation impingent on a step potential. Let the potential barrier be of height V, the motion of the particle be along the positive z, and the potential at $z = 0$. It's clear that in the region $z < 0$ we'll have the incident wave $\psi_I(x) = e^{ikz}$ with wave number $k = \frac{\sqrt{2mE}}{\hbar} = \frac{p_z}{\hbar}$, and the oppositely propagating reflected wave also with wave number k and $\psi_R = a_R e^{-ikz}$ where a_R is a complex coefficient. In the region $z > 0$ we will have the transmitted wave $\psi_T = a_T e^{ik_T z}$ with $k_T = \frac{\sqrt{2m(E-V)}}{\hbar}$, note that if $E < V$ k_T will be purely imaginary and the wave in the region $z > 0$ will be exponentially damped. We now match the wave function and its first derivative at $z = 0$ and we get following two simultaneous equations

$$\begin{aligned} a_T &= 1 + a_R \\ ik_T a_T &= ik - ika_R \end{aligned} \tag{3.148}$$

which are easily solved to obtain:

$$\begin{aligned} a_R &= \frac{k - k_T}{k + k_T} \\ a_T &= \frac{2k}{k + k_T} \end{aligned} \tag{3.149}$$

To calculate the probability of transmission/reflection recall from chapter 2 that the probability density and probability current density for the Schrödinger equation are given by $\rho = \psi^*\psi$ and $\vec{j} = \frac{\hbar}{2mi}(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*)$ respectively and the continuity equation is

$$\frac{\partial}{\partial t}(\rho) + \vec{\nabla} \cdot \vec{j} = 0 \tag{3.150}$$

this of course simply reduces to

$$\vec{\nabla} \cdot \vec{j} = 0 \tag{3.151}$$

for this steady state problem. Thus $\frac{\partial j_z}{\partial z} = 0$ holds and $j_z \mid_{z<0} = j_z \mid_{z>0}$. We expect the the incident and reflected waves to be oscillatory in the potential free $z < 0$ region in the $z > 0$ region however we must consider the possibilities $E > V$ and $E < V$ thus a real or an imaginary transmitted wave number k_T and we take care to insert the appropriate wave function for ψ_T

Calculating the current one easily gets $\frac{\hbar k}{m} - \mid a_R \mid^2 \frac{\hbar k}{m} = Re(k_T)\frac{\hbar}{m} \mid a_T \mid^2$, rearranging and dividing by $\frac{\hbar k}{m}$ we obtain

$$\frac{Re(k_T)}{k} \mid a_T \mid^2 + \mid a_R \mid^2 = 1$$

Note that we can define the ratios of the transmitted and reflected currents to the incident and define the probabilities of transmission and reflection respectively as $T = \frac{k_T}{k} \mid a_T \mid^2$ and $R = \mid a_R \mid^2$ respectively. Using a_T, a_R from 3.149 one obtains quite easily for when $E > V$

$$\frac{4kk_T}{(k + k_T)^2} + \frac{k - k_T}{(k + k_T)^2} = 1 \tag{3.152}$$

and so the probability of transmission and reflection are $T = \frac{4kk_T}{(k+k_T)^2}$ and $R = (\frac{k-k_T}{k+k_T})^2$ respectively. For $E < V$ and complex k_T the wave function ψ_T is a decaying exponential and its clear that the current is 0 in the region $z > 0$, in this case $R = 1$ and $T = 0$.

We now turn our attention to the Dirac equation and examine the same problem. It is convenient to do this in Hamiltonian form, and we assume an incident wave propagating along z, we add a potential barrier $V = e\Phi$, this is easy to include in the Hamiltonian form of the equation, we have for the two regions $z < 0$ and $z > 0$ the following three equations:

$$\begin{aligned} E\psi_I &= (c\vec{\alpha}\cdot\vec{p} + \beta mc^2))\psi_I \\ E\psi_R &= (c\vec{\alpha}\cdot\vec{-p} + \beta mc^2))\psi_R \\ (E - e\Phi)\psi_T &= (c\vec{\alpha}\cdot\vec{p'} + \beta mc^2))\psi_T \end{aligned} \tag{3.153}$$

where the subscripts $_T$ and $_I$ and $_R$ refer to transmitted, incident and reflected as in the Schrödinger equation example. We now replace $e\Phi$ by V and write down the following two dispersion relations in the two regions

$$\begin{aligned} E^2 &= c^2p^2 + m^2c^4 \\ (E - V)^2 &= c^2p_T^2 + m^2c^4 \end{aligned}$$

where we p is understood to mean a single component of momentum for this one dimensional problem an dT simply denotes that this relates to the transmitted wave. So now we have the wave numbers for the two regions

$$\begin{aligned} k &= \frac{\sqrt{E^2 - m^2c^4}}{c\hbar} \\ k_T &= \frac{\sqrt{(E - V)^2 - m^2c^4}}{c\hbar} \end{aligned}$$

its' easy to see that if $(E - V)^2 > m^2c^4 > 0$ we have a real k_T corresponding to a propagating wave in the region $z > 0$, and for $(E - V)^2 < m^2c^4$ we have an evanescent exponentially decaying wave. Now note that $(E - V)^2 - m^2c^4 = (E - V + mc^2)(E - V - mc^2)$ if both the factors are negative we have an oscillatory transmitted wave, as well as if both are positive. If one factor is negative and the other not then we have an exponential decay in the region $z > 0$. Thus we must consider *both possibilities* for a transmitted oscillatory wave-unlike in the case of a Schrödinger particle. To summarize

$$
\begin{aligned}
&E > V + mc^2, k_T = \frac{\sqrt{(E-V)^2 + mc^2}}{c\hbar} \\
&E < V - mc^2, k_T = \frac{\sqrt{(E-V)^2 + mc^2}}{c\hbar} \\
&V + mc^2 > E > V - mc^2, k_T = i\frac{\sqrt{(E-V)^2 + mc^2}}{c\hbar}
\end{aligned}
$$

With the wave numbers for the *three different cases* defined we now use equations 3.107 and write down incident, transmitted and reflected wave functions, where consider all helicity states. So we have the incident wave

$$\psi_I = \begin{pmatrix} 1 \\ 0 \\ \frac{c\hbar k}{E+mc^2} \\ 0 \end{pmatrix} e^{ik_z}$$

the transmitted wave

$$\psi_T = a_T \begin{pmatrix} 1 \\ 0 \\ \frac{c\hbar k_T}{E-V+mc^2} \\ 0 \end{pmatrix} e^{ik_T z} + b_T \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{c\hbar k_T}{E-V+mc^2} \end{pmatrix} e^{ik_T z} \qquad (3.154)$$

and the reflected wave

$$\psi_R = a_R \begin{pmatrix} 1 \\ 0 \\ \frac{-c\hbar k}{E+mc^2} \\ 0 \end{pmatrix} e^{-ikz} + b_T \begin{pmatrix} 0 \\ 1 \\ 0 \\ -\frac{c\hbar k}{E+mc^2} \end{pmatrix} e^{ikz} \qquad (3.155)$$

where the signs of k are set depending on the direction of propagation. Now the Dirac equation is linear in the space derivative and thus we need only match the wave functions themselves at the boundary and not the first derivatives. Since this is the result of linearizing a second order equation however, we must therefore match the several components. So applying $\psi_I + \psi_R = \psi_T$ and matching the components of the wave functions we arrive at the conditions:

$$
\begin{aligned}
&1 + a_R = a_T \\
&b_R = b_T \\
&\frac{c\hbar k}{E + mc^2} - a_R \frac{c\hbar k}{E + mc^2} = a_T \frac{c\hbar k_T}{E - V + mc^2} \\
&b_R \frac{c\hbar k}{E + mc^2} = -b_T \frac{c\hbar k_T}{E - V + mc^2}
\end{aligned}
\tag{3.156}
$$

Its easy to see that the second and third condition can only be met if $b_R = b_T = 0$ thus there is no change in the spin of the incident wave. We now define

$$
\alpha = \frac{k_T}{k} \frac{E + mc^2}{E - V + mc^2}
\tag{3.157}
$$

and solve for a_R and a_T

$$
\begin{aligned}
a_R &= \frac{1 - \alpha}{1 + \alpha} \\
a_T &= \frac{2}{1 + \alpha}
\end{aligned}
\tag{3.158}
$$

note that k_T must be treated as complex, since the transmitted wave can be oscillatory or evanescent and as with the Schrödinger case we define the probabilities of reflection and transmission to be the ratio of the reflected current to the incident, and the transmitted current to the incident. The probability current was defined in chapter 2 it is $c\psi^\dagger \alpha_3 \psi$ where we must insert the appropriate ψ. For the transmitted wave we calculate

$$
a_T a_T^* \left(1, 0, \frac{c\hbar k_T^*}{E - V + mc^2}, 0\right) \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{c\hbar k_T}{E - V + mc^2} \\ 0 \end{pmatrix}
\tag{3.159}
$$

whilst keeping in mind that k_T in particular can be real or imaginary. It's easy to do this for each of the incident reflected and transmitted waves. Dividing each of these fluxes by the incident flux we write down the fluxes normalized to the incident flux in the following form, using j and the appropriate subscripts

$$
\begin{aligned}
j_I &= 1 \\
j_R &= \mid \frac{1-\alpha}{1+\alpha} \mid^2 \\
j_T &= \mid \frac{2}{1+\alpha} \mid^2 \frac{Re(k_T)E + mc^2}{k(E - V + mc^2)} = \mid \frac{2}{1+\alpha} \mid^2 Re(\alpha)
\end{aligned}
$$

note that α (equation 3.157) can only be imaginary when k_T is imaginary, this happens when $V + mc^2 > E > V - mc^2$ as summarized in equations 3.154. If k_T is imaginary the behaviour of the currents is easy to understand the transmitted current is zero as is seen from j_T in equation 3.160, since with α imaginary

$$j_R = \mid \frac{1-\alpha}{1+\alpha} \mid^2 = \frac{(1-\alpha^*)(1-\alpha^*)}{(1+\alpha)(1+\alpha^*)} \mid^2$$

which is simply 1 equalling the incident current. Now for the interesting bit: if k_T is real and $E > V + mc^2$ we have an oscillatory wave function, all is well with $j_I = j_R + J_T$, however if we “turn up” the value of V then its easy to see that j_T changes corresponds to entering the region $z < 0$ where we shouldn’t have a transmitted current ! In addition to this the denominator in $\frac{2}{1+\alpha}$ now becomes less than 1.

Chapter 4

Propagators and scattering

In this chapter our goal is to develop a formalism to calculate quantum mechanical scattering amplitudes of spin $\frac{1}{2}$ particles described by the Dirac equation in the presence of an electromagnetic potential. Thus particles at $t' = -\infty$ are prepared in an incoming free wave state of an initial momentum and energy, interact with an electromagnetic potential and emerge after scattering in an outgoing free wave state of different momentum and energy at $t' = +\infty$ (reversed for the description of positrons as we shall see later in this chapter). The scattered particles are detected and our goal is to understand the nature and strength of the interaction.

In later chapters we'll assume that a beam of particles is represented by a flux–a probability per unit area per unit time. The amplitude described in the first paragraph can be used to calculate a cross-section or an effective area which when multiplied by the flux gives the total number of scattered particles. The differential cross section can then be expressed in terms of a differential of a scattering angle, momentum or energy we refer to it as a differential cross section. In this chapter we will not calculate cross sections but will come up with a formalism for deriving the probability *amplitudes* for scattering processes with the interaction represented by a weak potential. Since most potentials cannot be exactly solved for, this formalism will allow us to solve for the probability amplitudes perturbatively upto a desired order of accuracy. The perturbing four potential that we'll introduce, $A^\mu(x)$, can be due to the presence or motion of another charged particle or it can be the 4-vector potential representing a photon as we'll see in later chapters when several physical processes will be calculated.

At the heart of our scattering formalism lies the propagator or Green's function which relates a quantum mechanical wave function at one point in space and time to the wave function at another point in space and time.

Every point on the wave front acts a source for the wave at a later point of time in accordance with Huygen's principle [4]. The propagator in the absence of a potential, or the free propagator, will be calculated first. The propagator in the presence of a potential or the full propagator will be calculated as a perturbative series containing the free propagator and succesively higher powers of the potential. We will begin with the Schrödinger equation as an example-which defines the non-relativistic propagator and then move on to the Dirac equation and derive the relativistic propagator known as the Feynman propagator. We will also derive an expression for S matrix containing transition amplitudes for scattering from an initial to a final state as a perturbative series. This approach is covered in several text books [7] [9] [12] [9] and this book follows these approaches closely. At this point switch to a system of units explained and justified in chapter 1 section 1.15, with $\hbar = 1$ and $c = 1$.

4.1 The definition of the free and full non-relativistic propagator

To illustrate the method we will begin by deriving the Green's function or propagator for wave function solutions of the Schrödinger equation. This will be referred to alternately as the non-relativistic propagator or the Schrödinger propagator.

Let $\psi(x, t)$ represent such a wave function at a particular point in space and time x, t. According to Huygens principle, each point x acts as a source for the wave function at another point x' at a time $t' > t$. To obtain the wave at a later time t' all the contributions from the wave front at t must be summed up taking into account the particular differential equation that the wave satisfies. A very similar set of steps will be followed in later chapters for the relativistic propagagator or the propagator for the Dirac equation-also known as the Feynman propagator. We postulate the existence of a function, called a propagator or (Green's function) that relates ψ at (x, t) to ψ at (x', t') by the equation

$$\psi(x', t') = i \int d^3\vec{x} G(x', t'; x, t)\psi(x, t) \tag{4.1}$$

Its important to note that the Green's function in Eqn. 4.1 is assumed to give the propagated wave function in the presence of a perturbing potential. We will call Green's functions with this property the "full" green's function or propagator. We'll shortly introduce a Green's function or propagator for a free wave satsifying the potential free Schrödinger equation.

Initially we had stated that we wanted to consider a wave function in time that has evolved from one earlier in time and so we now change our notation slightly to take into account causality, *i.e.* that a propagated wave appears only at time $t' > t$ and write:

$$\psi(x', t') = i \int d^3\vec{x} G^+(x', t'; x, t)\psi(x, t) \tag{4.2}$$

where the superscript $^+$ tells us that the Green function in Eqn. 4.2 relates a wave function earlier in time to one later in time. We will *impose* this condition mathematically a little later. We also now define the propagator or Green's function $G_0^+(x', t'; x, t)$ for a free wave $\phi(x, t)$ in the absence of a potential

$$\phi(x', t') = i \int d^3\vec{x} G_0^+(x', t'; x, t)\phi(x, t) \tag{4.3}$$

where $\phi(x, t)$ etc are the free wave solutions to the Schrödinger equation. We now return to the wave function $\psi(x', t')$ which satisfies the Schrödinger equation in the presence of an external potential (note $\hbar = c = 1$)

$$i\frac{\partial}{\partial t'}\psi(x', t') = (-\frac{1}{2m}\nabla'^2 + V(x'))\psi(x', t') \tag{4.4}$$

which we will rewrite. We make the following change of notation for convenience before proceeding

$$\begin{aligned}
\hat{H}'_0 &= -\frac{1}{2m}\nabla'^2 \\
\hat{H}' &= -\frac{1}{2m}\nabla'^2 + V(x') \\
i\partial_{t'} &= \frac{i\partial}{\partial t'}
\end{aligned} \tag{4.5}$$

representing respectively the free Schrödinger Hamiltonian, the full Schrödinger Hamiltonian and $i\times$ the time derivative in terms of the primed co-ordinates where we observe the scattered wave. We now write the Schrödinger equation in the form

$$(i\partial_{t'} - \hat{H}')\psi(x', t') = 0 \tag{4.6}$$

and using the theta or Heaviside function ($\theta(t'-t) = 0$ for $t > t'$, $= 1$ for $t' > t$) we now formally impose the requirement that no scattering takes place unless $t' > t$ using the Heavyside or theta function and write down the propagators in the presence and absence of a potential

$$\begin{aligned}\theta(t'-t)\psi(x',t') &= i\int d^3\vec{x}G^+(x',t';x,t)\psi(x) \\ \theta(t'-t)\phi(x,t) &= i\int d^3\vec{x}G_0^+(x',t';x,t)\phi(x,t)\end{aligned} \tag{4.7}$$

We will now digress slightly to demonstrate how the Heaviside function can be represented analytically, this representation will allow us to find the differential equation satisfied by the Green's function or propagator.

4.2 Integral representation of the θ function

We now claim that the θ function's characteristics are represented by the following integral

$$\theta(t'-t) = \lim_{\epsilon\to 0} i\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-i\omega(t'-t)}}{\omega+i\epsilon}d\omega \tag{4.8}$$

The complex number ω in Eqn. 4.8 in terms of its real and imaginary parts (subscripts $_R$, $_I$) is

$$\omega = \omega_R + i\omega_I$$

the exponent in the integral Eqn. 4.8 is then: $i\omega_R(t'-t) + \omega_I(t'-t)$. So if $t'-t > 0$, then the exponential in the integral in Eqn. 4.8 is damped in the lower half complex plane, on the other hand if $t'-t < 0$ the exponent is damped in the upper half plane. The denominator of the integral Eqn. 4.8 contains a pole at $\omega = (0, -i\epsilon)$ (lower half plane).

A contour integral can now be defined as shown in Fig. 4.1 we take the contours C_1, C_2 (summarized as $C_{1,2}$) to mean a contour including the semi-circular part in the upper or lower half complex planes respectively. The integral over either of these contours is on the left hand side of Eqn. 4.9. Note that the integral over the contour in the upper half plane contains no pole and is zero by Cauchy's integral formula [2]. The contour in the lower half plane (traversed clockwise) is by Cauchy's integral formula $= 2\pi i(-1)(-1)\frac{1}{2\pi i}e^0 = 1$. Now the contour integral is the sum of the line integral along the real axis and the semi-circular part in Eqn. 4.9. If the radius ρ of either semi-circular part (denoted by $S_{1,2}$) is allowed to go to

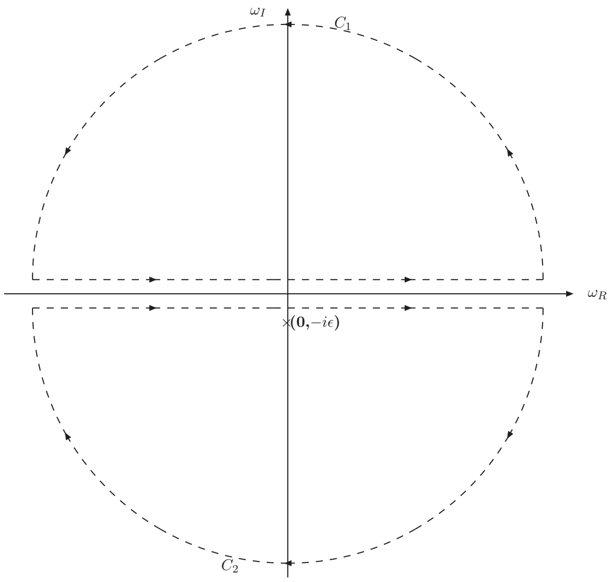

Fig. 4.1 Integration contour for representation of $\theta(t'-t)$ function (radius of $C_{1,2} \to \infty$) When $t'-t$ is positive the contour is closed including the pole in lower half plane where the exponent is damped yielding exactly 1 (see text). If $t'-t$ is negative then the exponent is damped in the upper half plane-the contour there encloses no poles and so the integral is zero replicating the behaviour of $\theta(t'-t)$.

∞ then it's contribution is zero due to the exponential damping and the integral along the real axis in Eqn. 4.8 is equal to the contour integral on the left. Thus the equality below holds for a contour closed in the upper-half plane if $t'-t<0$ and the lower half plane if $t'-t>0$, using Jordan's Lemma [2].

$$\begin{aligned} & i\oint_{C_{1,2}} \frac{1}{2\pi}\frac{e^{-i\omega(t'-t)}}{\omega+i\epsilon}d\omega = \\ & \lim_{\epsilon\to 0} i\frac{1}{2\pi}\Big(\int_{-\rho}^{\rho}\frac{e^{-i\omega(t'-t)}}{\omega+i\epsilon}d\omega_R + \int_S \frac{e^{-i\omega(t'-t)}}{\omega}d\omega\Big) \end{aligned} \tag{4.9}$$

By the behaviour of the contour integral in Eqn. 4.9 for the two cases of damping in the upper and lower half planes we conclude that the integral along the real axis- Eqn. 4.8 behaves like the θ function with argument $t'-t$ as $\rho \to \infty$.

4.3 The differential equation satisfied by the Green's function

We begin by noting that the partial derivative of the integral representation of the θ function (just verified in the previous section)

$$\lim_{\epsilon \to 0} i \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\omega_R(t'-t)} e^{-\omega_I(t'-t)}}{\omega + i\epsilon} d\omega$$

with respect to t' is simply $\int_{-\infty}^{\infty} e^{-i\omega(t'-t)} \frac{d\omega}{2\pi}$ which is simply one representation of the Dirac Delta function $\delta(t'-t)$ [2] and so:

$$\partial_{t'} \theta(t'-t) = \delta(t'-t) \tag{4.10}$$

We now operate on both sides of Eqn. 4.7 with the operator $i\partial_{t'} - \hat{H}'$ and obtain

$$\begin{aligned} &(i\partial_{t'} - \hat{H}')(\theta(t'-t)\psi(x')) \\ &= i \int d^3\vec{x}((i\partial_{t'} - \hat{H}')G^+(x',t';x,t))\psi(x) \end{aligned} \tag{4.11}$$

using the product rule, recognizing that $(i\partial_{t'} - \hat{H}')\psi(x') = 0$ and $i\partial_{t'}\theta(t'-t) = \delta(t'-t)$ (discussion preceding and including Eqn. 4.10) we obtain :

$$\delta(t'-t)\psi(x') = i \int d^3\vec{x}((i\partial_{t'} - \hat{H}')G^+(x',t';x,t))\psi(x) \tag{4.12}$$

Recognizing that the propagation of a wave under any conditions from one space-time point to another cannot depend on the co-ordinates themselves but only on their differences we replace $G^+(x',t';x,t)$ and $G_0^+(x',t';x,t)$ by $G^+(x'-x)$ and $G_0^+(x'-x)$ where the x' and x are now taken to include the time and position co-ordinates.

Now back to equation 4.12: Note that the unprimed variables are unaffected by the operator $i\partial_{t'} - \hat{H}'$ and using the properties of the Dirac delta function we replace $\psi(x')$ by $\int d^3\vec{x}\delta^3(x'-x)\psi(x)$ and get

$$\begin{aligned} &i\delta(t'-t) \int d^3\vec{x}\psi(x)\delta^3(\vec{x}'-\vec{x}) = \\ &i \int d^3\vec{x}(i\partial_{t'} - \hat{H}')(G^+(x'-x))\psi(x) \end{aligned} \tag{4.13}$$

and can now simply now remove $i\int d^3\vec{x}\psi(x)$ on both sides of Eqn. 4.13 and finally obtain the differential equation that must be satisfied by the retarded propagator: $G^+(x';x)$:

$$(i\partial_{t'} - \hat{H}')G^+(x'-x) = \delta^3(\vec{x}'-\vec{x})\delta(t'-t) = \delta^4(x-x') \tag{4.14}$$

We can take the operator $\hat{O}' = i\partial_{t'} - \hat{H}'$ to represent the difference between the time derivative and the Hamiltonian for any wave equation, we then summarize Eqn. 4.14

$$\hat{O}'G^+(x'-x) = \delta^4(x'-x) \tag{4.15}$$

where $\hat{O}' = i\partial_{t'} - \hat{H}'$ with $\hat{H}'$ representing the Schrödinger equation Hamiltonian (including the potential V), the operator $\hat{O}'$ can be generalized for any wave equation using the time derivative operator and relevant Hamiltonian. We will heretofore refer to the propagator in the presence of a potential as the "full propagator" $(G^+(x'-x))$ and the propagator in the absence of a potential as the "free propagator"$(G_0^+(x'-x))$. The following is the equation satisfied by the free propagator

$$\hat{O}'_0 G_0^+(x'-x) = \delta^4(x'-x) \tag{4.16}$$

where $\hat{O}'_0 = i\partial_{t'} - \hat{H}'_0$. As mentioned earlier the scattering potential is assumed to not have an exact solution, we will therefore develop an approximate solution, to do this we will first solve Eqn. 4.16. We do this by using Fourier transforms, just like any other function in the universe we can write the Greens function as a Fourier transform, thus

$$G_0^+(x'-x) = \frac{\int d^3\vec{p}d\omega}{(2\pi)^4} e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} \times e^{-i\omega(t'-t)} G_0^+(p,\omega) \tag{4.17}$$

where $\vec{p}$ and ω represent the momentum and energy which have dimensions of $\frac{1}{\text{SPACE}}$ and $\frac{1}{\text{TIME}}$ and so the exponent in Eqn. 4.17 is unitless. Once we determine $G_0^+(p,\omega)$ and we will have our free propagator in co-ordinate space (as a function of $x'-x$, as originally written). We now rewrite the Dirac-Delta function on the right hand side of Eqns. 4.14 or 4.15 simply by using the same variables in the Fourier transform, so equation (4.14) becomes:

$$\left(i\partial_{t'} + \frac{1}{2m}\nabla'^2\right)\frac{\int d^3\vec{p}d\omega}{(2\pi)^4}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})-i\omega(t'-t)}G_0^+(p,\omega)$$
$$= \frac{\int d^3\vec{p}d\omega}{(2\pi)^4}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})-i\omega(t'-t)} = \delta^4(x'-x) \tag{4.18}$$

where the reader will recognize $\frac{\int d^3\vec{p}d\omega}{(2\pi)^4}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})-i\omega(t'-t)}$ as the 4-dimensional Dirac Delta function [?].

We now operate $i\partial_{t'} + \frac{1}{2m}\nabla'^2$ on the right and remove the integral over $\int \frac{d^3\vec{p}d\omega}{(2\pi)^4}$ and the exponential on both sides:

$$(\omega - \frac{\vec{p}\cdot\vec{p}}{2m})G_0^+(p,\omega) = 1 \tag{4.19}$$

and so

$$G_0^+(p,\omega) = \frac{1}{\omega - \frac{\vec{p}\cdot\vec{p}}{2m}} \tag{4.20}$$

this is the "momentum space" Greens function or propagator, using Eqn. 4.17 we now write down the propagator in co-ordinate space:

$$G_0^+(x'-x) = \int \frac{d^3\vec{p}}{(2\pi)^4}d\omega \frac{e^{i\vec{p}\cdot(\vec{x}'-\vec{x})-i\omega(t'-t)}}{\omega - \frac{\vec{p}\cdot\vec{p}}{2m}} \tag{4.21}$$

Splitting up the integral in Eqn. 4.21 as a product of the time and space parts:

$$G_0^+(x'-x) = \int \frac{d^3\vec{p}}{(2\pi)^3}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{-i\omega(t'-t)}}{\omega - \frac{\vec{p}\cdot\vec{p}}{2m}} \tag{4.22}$$

rewriting the time integral in Eqn. 4.22 we obtain :

$$G_0^+(x'-x) = \int \frac{d^3\vec{p}}{(2\pi)^3}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})} \times \lim_{\epsilon\to 0}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{-i\omega(t'-t)}}{\omega - \frac{\vec{p}\cdot\vec{p}}{2m} + i\epsilon} \tag{4.23}$$

Recall that in the section 4.2 we had verified that

$$\lim_{\epsilon\to 0}\frac{-1}{2\pi i}\int_{-\infty}^{\infty}\frac{e^{i\omega(t'-t)}}{\omega + i\epsilon}d\omega = \theta(t'-t)$$

Thus Eqn. 4.23 can be rewritten :

$$G_0^+(x'-x) = -i\theta(t'-t)\int \frac{d^3\vec{p}}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}'-\vec{x}) - \frac{i\vec{p}\cdot\vec{p}}{2m}(t'-t)} \tag{4.24}$$

At this point we can identify the free wave solutions to the Scrödinger equation with momentum $\vec{p}$ in the integrand of Eqn. 4.24. These are simply the plane waves [11]

$$\phi(x') = \frac{e^{i\vec{p}\cdot\vec{x}' - \frac{i\vec{p}\cdot\vec{p}t'}{2m}}}{(2\pi)^{3/2}}, \phi^*(x) = \frac{e^{-i\vec{p}\cdot\vec{x} + \frac{i\vec{p}\cdot\vec{p}t}{2m}}}{(2\pi)^{3/2}}$$

Where the energy is $\omega = \frac{\vec{p}\cdot\vec{p}}{2m}$. With these solutions at hand we can write down the Green's function or propagator for the Schrödinger equation in terms of the free wave solutions

$$G_0^+(x'-x) = -i\theta(t'-t)\int d^3\vec{p}\phi(x',t')\phi^*(x,t) \tag{4.25}$$

We will now verify the behaviour of the propagator as advertised in Eqn. 4.7 Consider a free solution to the Schrödinger equation of momentum $\vec{p}'$ and energy $\omega' = \frac{\vec{p}'\cdot\vec{p}'}{2m}$ this is of course $\phi(x,p',\omega') = \frac{e^{i\vec{p}'\cdot x - i\omega' t}}{(2\pi)^{\frac{3}{2}}}$. We can now simply evaluate the right hand side of Eqn. 4.7 using the definition in Eqn. 4.25 with ϕ instead of ψ, and using the subscripts p and p' to denote the different momenta

$$\begin{aligned}
i\int d^3\vec{x}G_0^+(x'-x)\phi(x,t) &= -i(i)\theta(t'-t)\iint d^3\vec{p}d^3\vec{x}\phi_p(x',t')\phi_p^*(x,t)\phi_{p'}(x,t)\\
&= \theta(t'-t)\iint d^3\vec{p}\phi_p(x',t')\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-i\vec{p}\cdot\vec{x}+i\omega t}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{i\vec{p}'\cdot\vec{x}-i\omega' t}\\
&= \theta(t'-t)\iint d^3\vec{p}\phi_p(x',t')d^3\vec{x}\frac{1}{(2\pi)^3}e^{-i(\vec{p}-\vec{p}')\cdot\vec{x}-i(\omega'-\omega)t}\\
&= \theta(t'-t)\int d^3\vec{p}\phi_p(x',t')\delta^3(\vec{p}'-\vec{p})e^{i(\omega'-\omega)t} = \theta(t'-t)\phi_p(x',t')
\end{aligned} \tag{4.26}$$

in Eqn. 4.26 the definition of the Dirac delta function has been used. Integrating over $d^3\vec{p}$ sets $\vec{p}' = \vec{p}$ and $\omega' = \omega = \frac{\vec{p}\cdot\vec{p}}{2m}$ and we are left with $\theta(t'-t)\phi(x',t')$ where the subscripts p, p' distinguishing the two momenta has been dropped. Its easy to see that the defining equation for the free propagator Eqn. 4.7 has been satisfied. It is straightforward to write down a free propagator for a *discrete* set of solutions as well

$$G_0^+(x'-x) = -i\theta(t'-t)\sum_n \phi_n(x',t')\phi_n^*(x,t) \tag{4.27}$$

to verify this we multiply it on the right first by $\phi_m(x,t)$ integrate over $d^3\vec{x}$ to obtain

$$\begin{aligned}
&\int G_0^+(x'-x)\phi_m(x,t)d^3\vec{x} = \\
&-i\theta(t'-t)\sum_n \int \phi_n(x',t')\phi_n^*(x,t)\phi_m(x,t)d^3\vec{x} \\
&i\int G_0^+(x'-x)\phi_m(x,t)d^3\vec{x} = -i\theta(t'-t)\sum_n \phi_n(x',t')\delta_{mn} = \theta(t'-t)\phi_m(x',t')
\end{aligned} \tag{4.28}$$

the last line of Eqn. 4.28 is the definition of the propagator. We can also determine the effect of this propagator on $\phi_m^*(x',t')$, here we integrate over $d^3\vec{x}'$ and it is straightforward to see that:

$$\begin{aligned}
&\int G_0^+(x'-x)\phi_m^*(x',t')d^3\vec{x}' = \\
&i\theta(t'-t)\int \sum_n \phi_m^*(x',t')\phi_n(x',t')\phi_n^*(x,t)d^3\vec{x}' = \\
&i\int G_0^+(x'-x)\phi_m^*(x',t')d^3\vec{x}' = \theta(t'-t)\phi_m^*(x,t)
\end{aligned} \tag{4.29}$$

Eqn. 4.29 requires a bit of attention: t' is a later point of time than t, thus 4.29 is representing the propagation of the $\phi^*(x',t')$ backward in time to t.

By using ψ to denote a wave function that is an exact solution in the presence of a potential we can write down the following expressions for the propagator in the presence of a potential-the exact propagator

$$G^+(x'-x) = -i\theta(t'-t)\sum_n \psi_m^*(x',t')\psi_n(x',t') \tag{4.30}$$

since we are developing a method for potentials such that Schrödinger equation does not have an exact solution (*i.e.* $\psi_m(x,t)$ is an approximate solution) and so we don't have an exact solution for $G^+(x'-x)$, we need to ahave all solutions of the Schrödinger equation to do this, therefore we will solve for $G^+(x'-x)$ as an approximate, perturbative series in the potential V and $G_0^+(x'-x)$.

4.4 The propagator in the presence of a potential. The "full" propagator.

In this section we will solve for $G^+(x'-x)$ (the full propagator, containing the effects of the potential) as an approximate integral series in $G_0(x'-x)$ and $V(x')$. For the series to converge the assumption is that V is small. Recalling Eqn. 4.14

$$(i\frac{\partial}{\partial t'} - \hat{H}')G^+(x';x) = \delta^4(x-x') \tag{4.31}$$

where the operator $\hat{H}' = -\frac{1}{2m}\nabla^{2\prime} + V(x')$. Using this Eqn. 4.31 becomes

$$(i\partial_{t'} - \hat{H}_0')G^+(x'-x) = \delta^4(x'-x) + V(x')G^+(x'-x) \tag{4.32}$$

using the properties of the Dirac delta function we can rewrite the right hand side of Eqn. 4.32

$$\begin{aligned}(i\partial_{t'} - \hat{H}_0')G^+(x'-x) &= \delta^4(x'-x) \\ &\quad + \int d^4x_n \delta^4(x'-x_n)V(x_n)G^+(x_n-x)\end{aligned} \tag{4.33}$$

where x_n is an arbitrary dummy variable comprising three space and one time co-ordinate. We now use the definition of the free propagator appropriately

$$(i\partial_{t'} - \hat{H}_0')G_0^+(x'-x_n) = \delta^4(x-x_n) \tag{4.34}$$

and replace the Dirac delta functions on the right hand side of Eqn.4.33 :

$$\begin{aligned}(i\partial_{t'} - \hat{H}_0')G^+(x'-x) &= (i\partial_{t'} - \hat{H}_0')(G_0^+(x'-x) \\ &\quad + \int d^4x_n G_0^+(x'-x_n)V(x_n)G^+(x_n-x))\end{aligned} \tag{4.35}$$

If $V(x')$ were $= 0$ then $G^+(x'-x)$ would have to be $= G_0(x'-x)$ by definition, so we can remove this operator $(i\partial_{t'} - \hat{H}_0')$ from both sides and obtain:

$$\begin{aligned}G^+(x'-x) &= G_0^+(x'-x) \\ &\quad + \int d^4x_n G_0^+(x'-x_n)V(x_n)G^+(x_n-x)\end{aligned} \tag{4.36}$$

Equation 4.36 is an integral equation for $G^+(x'-x)$ in terms of the free propagator $G_0^+(x'-x)$ and the potential V and is known as the Lippman-Schwinger equation. Since we know that G^+ cannot be solved for exactly we expect that we will have to iterate this equation to a desired order in V. Using Eqn. 4.36 to write $G^+(x_n-x)$ using the Lippman-Schwinger equation itself with x_n in place of x' and introducing another dummy variable x_{n-1} we obtain:

$$G^+(x_n-x) = G_0^+(x_n-x) + \int d^4x_{n-1} G_0^+(x_n - x_{n-1}) V(x_{n-1}) G^+(x_{n-1}-x)$$

and so Eqn. 4.36 becomes

$$\begin{aligned} G^+(x'-x) = G_0^+(x'-x) &+ \int d^4x_n G_0^+(x'-x_n) V(x_n) G^+(x_n-x) \\ &+ \iint d^4x_n d^4x_{n-1} G_0^+(x'-x_n) V(x_n) G_0^+(x_n-x_{n-1}) \times \\ &V(x_{n-1}) G^+(x_{n-1}-x) \end{aligned} \tag{4.37}$$

this iterative procedure can be repeated again, each time substituting for G^+ on the right hand side, using another dummy variable. We set $n=1$ in the first integration, $n=2$ in the second integration on the right, leave it to the reader to calculate the next iterative term, set $n=3$, and obtain the series:

$$\begin{aligned} &G^+(x'-x) = G_0^+(x'-x) + \int d^4x_1 G_0^+(x'-x_1) V(x_1) G_0^+(x_1;x) \\ &+ \iint d^4x_2 d^4x_1 G_0^+(x'-x_2) V(x_2) G_0^+(x_2-x_1) V(x_1) G_0^+(x_1-x) \\ &+ \iiint d^4x_3 d^4x_2 d^4x_1 G_0^+(x'-x_3) V(x_3) G_0^+(x_3-x_2) \\ &\times V(x_2) G_0^+(x_2-x_1) V(x_1) G_0^+(x_1-x) + \cdots \end{aligned} \tag{4.38}$$

We assume that the series will converge due to the smallness of V. The reader is reminded that in the first term on the right hand side by definition of the retarded propagator $x'^0 = t' > x^0 = t$, in the second term $t' > t_1 > t$ in the third $t' > t_2 > t_1 > t$ term etc, of course it should be obvious that $t' > t_n > t_{n-1} \cdots > t_3 > t_2 > t_1 > t$. The propagator can be seen to be expressed as the sum of the free propagator and a series of successive scatterings at consecutive times at distinct space time points with free propagation between them. The series can be calculated to the desired degree of accuracy.

4.5 Using the propagator to calculate probability amplitudes: The S matrix

We will now use the Green's function or propagator approach to calculate an expression for the S matrix that contains the scattering amplitudes from free initial to a free final state of specified momenta. As stated we'll assume that at $t' = \infty$ and $t = -\infty$ the waves are free so the potential must satisfy

$$\lim_{t'\to\infty} V(x', t') = 0, \quad \lim_{t\to-\infty} V(x, t) = 0$$

with the corresponding conditions on the initial and final states of the particle

$$\lim_{t'\to\infty} \psi(x', t') = \phi_f(x', t'), \quad \lim_{t\to-\infty} \psi(x, t) = \phi_i(x, t) \tag{4.39}$$

where the $\phi_{i,f}$ represent the free initial and final states of momenta $\vec{p_i}$ and $\vec{p_f}$ respectively. At the end of the day we are interested in calculating quantities that give us the probability amplitude for a transition from one state to the next. We now define the S matrix:

$$S_{fi} =< \psi_f(x', t') \mid \psi(x', t' >= \int d^3\vec{x}'\phi_f^*(x', t')\psi(x', t') \tag{4.40}$$

Where $\psi(x', t')$ defines the *full* wave propagated to x', t' and the limit $t' \to \infty$ is understood. This wave at $t = -\infty$ is equal to the initial incoming free wave with momentum $\vec{p_i}$ hence the subscript $_i$ in S_{fi}. Taking Eqn. 4.40 we can use the definition of the propagator and Eqn. 4.40 and write

$$\begin{aligned} S_{fi} &= \lim_{t'\to\infty} \int d^3\vec{x}'\phi_f^*(x', t')\psi(x', t') \\ &= \lim_{t'\to\infty} \lim_{t\to-\infty} \int d^3\vec{x}'\phi_f^*(x', t') i \int d^3x G^+(x' - x)\phi_i(x, t) \end{aligned} \tag{4.41}$$

with $\phi_i(x, t)$ representing the initial free plane wave. Inserting the derived series for $G^+(x' - x)$ in terms of V and G_0^+ from Eqn. 4.38 we obtain

$$\begin{aligned} S_{fi} = \lim_{t'\to\infty} \lim_{t\to-\infty} &\int d^3\vec{x}'\phi_f^*(x', t')(i \int d^3x (G_0^+(x' - x) \\ &+ \int d^4x_1 G_0^+(x' - x_1)V(x_1)G_0^+(x_1 - x) \\ &+ \iint d^4x_1 d^4x_2 G_0^+(x' - x_2)V(x_2)G_0^+(x_2 - x_1)V(x_1) \\ &\times G_0^+(x_1 - x))\phi_i(x, t)) \end{aligned} \tag{4.42}$$

Let's look at the first term in equation 4.42

$$i \iint d^3\vec{x}' d^3\vec{x} \phi_f^*(x') G_0^+(x'-x)\phi_i(x) \tag{4.43}$$

we can integrate over x' by using Eqn. 4.29 to obtain

$$\int d^3\vec{x}\phi_f^*(x)\phi_i(x)$$

since we are the $\phi_{f,i}$ are free particle solutions we can easily write the first term as

$$\int d^3\vec{x}\phi_f^*(x)\phi_i(x) = \int \frac{d^3\vec{x}}{(2\pi)^3} e^{-i(\vec{p_f}+\vec{p_i})\cdot x} = \delta^3(\vec{p}_f - \vec{p}_i)$$

by the definition of the Dirac delta function. Now we look at the second term in the series

$$i \iiint d^3\vec{x} d^4x_1 d^3\vec{x}' \phi_f^*(x') G_0^+(x'-x_1) V(x_1)\phi_i(x) G_0^+(x_1-x)$$

once again we integrate all terms that are functions of x' over $d^3\vec{x}'$, using equation 4.29 to obtain

$$\iint d^3\vec{x} d^4x_1 \phi_f^*(x_1) V(x_1) G_0^+(x_1-x)\phi_i(x)$$

rewriting this as

$$(-i)(i)\iint d^3\vec{x} d^4x_1 \phi_f^*(x_1) V(x_1) G_0^+(x_1-x)\phi_i(x)$$

we simply integrate using the definition of the propagator (hence we have inserted the extra i) using equation 4.28 to obtain the second term in this series:

$$(-i)\int d^4x_1 \phi_f^*(x_1) V(x_1)\phi_i(x_1) \tag{4.44}$$

It should be clear now that for every n^{th} term in the series we can integrate every

$$i\int d^3x' \phi_f^*(x') G_0^+(x'-x_n) \text{ and } \int d^4x G_0^+(x_1-x)\phi_i(x,t)$$

to lose the two three dimensional integrals over $d^3\vec{x}'$ and $d^3\vec{x}$ and obtain the series:

$$
\begin{aligned}
S_{fi} &= \delta^3(\vec{p}_f - \vec{p}_i) - i \int d^4x_1 \phi_f^*(x_1) V(x_1) \phi_i(x_1) \\
&- i \iint d^4x_1 d^4x_2 \phi_f^*(x_2) V(x_2) G_0^+(x_2 - x_1) V(x_1) \phi_i(x_1) \\
&- i \iiint d^4x_1 d^4x_2 d^4x_3 \phi_f^*(x_3) V(x_3) G_0^+(x_3 - x_2) V(x_2) \\
&\times G_0^+(x_2 - x_1) \phi_i(x_1) + \cdots
\end{aligned}
\tag{4.45}
$$

Examining Eqn. 4.45 we can see quite clearly that the first term on the right hand side is zero unless $\vec{p}_f = \vec{p}_i$ or if no scattering takes place. The probability amplitude for scattering is therefore represented as an approximation as a sum of amplitudes series of several scatterings to succesively higher order in V.

4.6 The relativistic (Feynman) propagator

In this section we will calculate the propagator for a Dirac particle in the presence of an electromagnetic field. The following convention is adopted for relativistic propagator or "Feynman propagator". We write down

$$
\begin{aligned}
&(i\,\not{\partial}' - e\,\not{A}(x') - m)_{\alpha\sigma} S_{F,\sigma\beta}(x' - x) = \delta^4(x' - x)\delta_{\alpha\beta} \\
&(i\,\not{\partial}' - m)_{\alpha\sigma} S^0_{F,\sigma\beta}(x' - x) = \delta^4(x' - x)\delta_{\alpha\beta}
\end{aligned}
\tag{4.46}
$$

The equations 4.46 are written in the spirit of equations 4.31 and 4.34. with the first equation defining the full propagator that takes into account the effects of a perturbing electro-magnetic potential and the second the free propagator. There are however key differences:

1. The full and free Feynman propagators $S_{F,\sigma\beta}(x'-x)$ and $S^0_{F,\sigma\beta}(x'-x)$ are 4×4 matrices which must act on a spinor wave-function-we will suppress the indices for convenience later on however.
2. Recall the operator $\hat{O}'$ defined for the Schrödinger equation in Eqns. 4.15 which includes the Hamiltonian and time derivative. The operator $i\,\not{\partial}' - e\,\not{A}(x')$ in equation 4.46 is in $\hat{O} = i\partial_{t'} - \hat{H'_{DIRAC}}$ multiplied by γ^0 where $\hat{H'_{DIRAC}}$ is the Dirac equation Hamiltonian first encountered in Chapter 2 (**??**). This factor of γ^0 missing in the defining Eqn. 4.46 will be recovered when defining the action of the propagator on the wave function later in Eqn. 4.66.

We will find $S_{F,\sigma\beta}(x'-x)$ defined in equation 4.46 by first finding $S^0_{F,\sigma\beta}(x'-x)$ and solving iteratively as we did for the Schrödinger equation. We begin following our previous technique by expressing $S^0_{F,\alpha\beta}(x'-x)$ as a Fourier transform

$$S^0_{F,\alpha\beta}(x'-x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)} S^0_{F,\alpha\beta}(p) \tag{4.47}$$

we now seek $S^0_{F,\alpha,\beta}(p)$ and afterward will develop a series solution for $S_{F,\alpha,\beta}(p)$ as well. We will for convenience's sake drop the matrix indices from now on. Demanding that the Fourier transform satisfy the conditions on the free propagator defined in Eqns. 4.46 and using $\delta^4(x'-x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)}$ for the Dirac delta function on the right hand side we have:

$$(i\,\partial\!\!\!/' - m) \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)} S^0_F(p) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)}$$

where the operator $(i\,\partial\!\!\!/' - m)$ operates on the exponential

$$\int \frac{d^4p}{(2\pi)^4} (p\!\!\!/ - m) e^{-ip\cdot(x'-x)} S^0_F(p) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x'-x)}$$

So we see that

$$(p\!\!\!/ - m) S^0_F(p) = 1 \tag{4.48}$$

where the reader is reminded that Eqn. 4.48 is a matrix equation and short-hand has been adopted-the 1 on the right hand side is in fact the identity matrix $\mathbb{I}$. Multiplying Eqn. 4.48 on the left with $p\!\!\!/ + m$ we obtain

$$(p\!\!\!/ + m)(p\!\!\!/ - m) S^0_F(p) = (p\!\!\!/ + m)$$

this is easily simpified as follows

$$p\!\!\!/\, p\!\!\!/ + m\, p\!\!\!/ - m\, p\!\!\!/ - m^2 = p\!\!\!/\, p\!\!\!/ - m^2$$

considering $p\!\!\!/\, p\!\!\!/$ first:

$$\begin{aligned}
p\!\!\!/\, p\!\!\!/ &= \gamma^\mu \gamma^\nu p_\mu p_\nu = (2g^{\mu\nu} - \gamma^\nu\gamma^\mu) p_\mu p_\nu \\
p\!\!\!/\, p\!\!\!/ &= 2g^{\mu\nu} p_\mu p_\nu - \gamma^\nu \gamma^\mu p_\nu p_\mu \\
p\!\!\!/\, p\!\!\!/ &= 2p\cdot p - p\!\!\!/\, p\!\!\!/ \\
p\!\!\!/\, p\!\!\!/ &= p\cdot p
\end{aligned} \tag{4.49}$$

We can see from Eqns. 4.49 that $p\!\!\!/\, p\!\!\!/ - m^2 = p\cdot p - m^2 = p^2 - m^2$ (multiplied by a suppressed identity matrix) and we can write

$$
\begin{aligned}
(p^2 - m^2) S_F^0(p) &= (\not{p} + m) \\
S_F^0(p) &= \frac{\not{p} + m}{p^2 - m^2}
\end{aligned}
\tag{4.50}
$$

The propagator in co-ordinate space is then using Eqn. 4.50

$$
S_F^0(x' - x) = \int \frac{d^4 p}{(2\pi)^4} \frac{\not{p} + m}{p^2 - m^2} e^{-ip\cdot(x'-x)} \tag{4.51}
$$

Note that the integral is over *all* momenta and not just those momenta that satisfy the relation $E^2 - \vec{p}\cdot\vec{p} - m^2 = 0$. (NEED TO TALK TO A THEORIST about virtual particles), this is related to the idea of "virtual particles"-in any scattering process we only ever observe the incoming and outgoing free particle states and not the propagation itself, the incoming and outgoing free states have masses that do satisfy $E^2 - \vec{p}\cdot\vec{p} - m^2 = 0$.

At this point we note that we could perform the integral over the four momenta in Eqn. 4.51 and get a closed function in terms of m, x', x, this is very well described in several texts [9] [15]. To make a long story short the propagator in co-ordinate space displays oscillatory behaviour in the forward and backward light cones for $E > 0$ and $E < 0$ solutions respectively and a damped exponential of opposite signs for each solution in the space-like region, this will not be described in detail here, and the reader is referred to the previous two excellent references for an insightful analysis. We will however *rewrite* the propagator in a more *useful form* that will allow us to interpret the negative energy and positve energy solutions in meaningful manner that provides a description of positrons and electrons, e^+, e^-. We now rewrite the propagator in the following manner:

$$
\begin{aligned}
S_F^0(x' - x) &= \lim_{\delta\to 0} \int \frac{d^4 p}{(2\pi)^4} \frac{\not{p} + m}{p^2 - m^2 + i\delta} e^{-ip\cdot(x'-x)} \\
&= \lim_{\delta\to 0} \int \frac{d^4 p}{(2\pi)^4} \frac{(\not{p} + m) e^{-ip\cdot(x'-x)}}{(p_0 - \sqrt{\vec{p}\cdot\vec{p} + m^2 - i\delta})(p_0 + \sqrt{\vec{p}\cdot\vec{p} + m^2 - i\delta})}
\end{aligned}
\tag{4.52}
$$

Taylor expanding the square root in the denominator of Eqn. 4.52 to first order we see that

$$
\sqrt{\vec{p}\cdot\vec{p} + m^2 - i\delta} \approx \sqrt{\vec{p}\cdot\vec{p} + m^2}(1 - i\frac{1}{2(\vec{p}\cdot\vec{p} + m^2)}\delta)
$$

we denote $\frac{1}{2(\vec{p}\cdot\vec{p}+m^2)}\delta$ by ϵ and rewrite the propagator in co-ordinate space broken up into an integral over just p_0 and $d^3\vec{p}$ in the following way:

$$
S_F^0(x'-x) = \lim_{\epsilon\to 0}\iint_{-\infty}^{\infty}\frac{d^3\vec{p}}{(2\pi)^3}\frac{dp_0}{(2\pi)}\frac{e^{-ip_0(t'-t)}e^{+i\vec{p}\cdot(\vec{x}'-\vec{x})}}{((p_0-E_p)+i\epsilon)((p_0+E_p)-i\epsilon)} \tag{4.53}
$$

where we have used $E_p = \sqrt{\vec{p}\cdot\vec{p}+m^2}$. Note that the integral on the right hand side of Eqn. 4.53 contains two poles at $p_0 = \mp E_p \pm i\epsilon = \mp\sqrt{\vec{p}\cdot\vec{p}+m^2} \pm i\epsilon$, the limits of integration along real p_0 (p_{0R}) range from and over each component of $\vec{p}$ range from $-\infty$ to $+\infty$. The reader is now asked to consider the following integral which contains a contour integral over dp_0 and a 3 dimensional integral over $d^3\vec{p}$ in which E_p denotes $\sqrt{\vec{p}\cdot\vec{p}+m^2}$

$$
\int\frac{d^3\vec{p}}{(2\pi)^3}\oint_{C_{1,2}}\frac{dp_0}{(2\pi)}\frac{(\not{p}+m)e^{-ip_0(t'-t)+i\vec{p}\cdot(\vec{x}'-\vec{x})}}{(p_0-E_p+i\epsilon)(p_0+E_p-i\epsilon)} \tag{4.54}
$$

the contour integral Eqn. 4.54 is only non zero when the contours $C_{1,2}$ enclose one of the two poles

$$
p_0 = \mp E_p \pm i\epsilon = \mp\sqrt{\vec{p}\cdot\vec{p}+m^2} \pm i\epsilon
$$

. The contours labelled $C_{1,2}$ in Eqn. 4.54 are shown in Fig. 4.2. The contours consist of a semi-circular part and an integral along the real axis with arbitrary limits $[-\rho,+\rho]$. The contour integral in Eqn. 4.54 can be written split up into parts along the real axis and the semicircular paths given by $S_{1,2}$ in a manner analogous to Eqn. 4.9

$$
\begin{aligned}
\int\frac{d^3\vec{p}}{(2\pi)^3}\oint_{C_{1,2}}\frac{dp_0}{(2\pi)}\frac{(\not{p}+m)e^{-ip_0(t'-t)+i\vec{p}\cdot(\vec{x}'-\vec{x})}}{(p_0-E_p+i\epsilon)(p_0+E_p-i\epsilon)} &= \\
\int\frac{d^3\vec{p}}{(2\pi)^3}(\int_{-\rho}^{\rho}\frac{dp_{0R}}{(2\pi)}\frac{\not{p}+m}{p^2-m^2+i\delta}e^{-ip\cdot(x'-x)}&+ \\
\int_{S_{1,2}}\frac{dp_0}{(2\pi)}\frac{\not{p}+m}{p^2-m^2+i\delta}e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}e^{-ip_{0I}(t'-t)})&
\end{aligned} \tag{4.55}
$$

and $S_{1,2}$ depicting the semi-circular parts of the contours (C_1, C_2). Note that if we let $\rho\to\infty$ then the first integral in the brackets on the right hand side in Eqn. 4.55 combined with the integral over dp_{0R} simply becomes $\int\frac{d^4p}{(2\pi)^4}\frac{\not{p}+m}{p^2-m^2}e^{-ip\cdot(x'-x)}$ which is $S_F^0(x'-x)$ and if the second integral in Eqn. 4.55 is damped along either semi-circular path then the sum of the two integrals by Jordan's Lemma [2] equals to the contour integral in Eqn. 4.54.

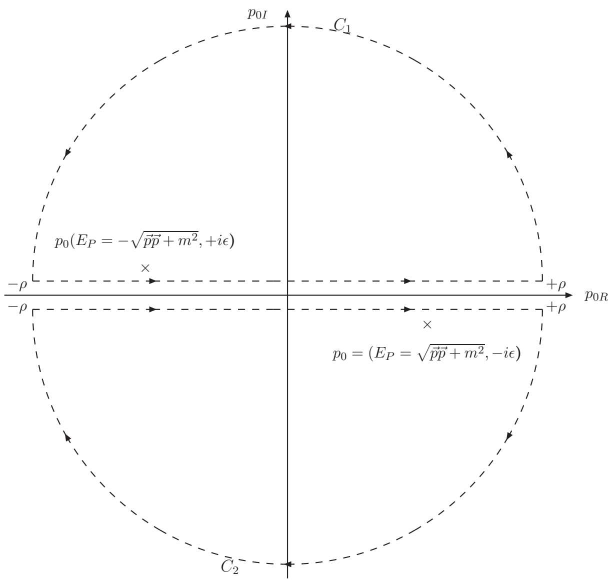

Fig. 4.2 Integration contours for the Feynman propagator (radius of $C_{1,2} \to \infty$). The integral over p_0 in Eqn 4.54 represents the contour integral over C_1 enclosing the pole at $-\sqrt{\vec{p}\cdot\vec{p}+m^2}+i\epsilon$ if the exponential in Eqn 4.55 is damped in the upper half plane and equals the contour over C_2 enclosing the pole at $\sqrt{\vec{p}\cdot\vec{p}+m^2}-i\epsilon$ if the damping is in the lower half plane. The reader can verify by writing $p_0 = p_{0R} \pm ip_{0I}$ that these cases represent $(t'-t) < 0$ and $(t'-t) > 0$ respectively.

It is easy to see that if $t'-t > 0$ then the second integral is exponentially damped in the lower half plane and we can close the contour C_1 enclosing the pole at $p_0 = \sqrt{\vec{p}\cdot\vec{p}+m^2}-i\epsilon$. If $t'-t < 0$ then the damping occurs in the upper half plane and the contour can be completed enclosing the pole at $p_0 = -\sqrt{\vec{p}\cdot\vec{p}+m^2}+i\epsilon$ in the upper half plane. Denoting $E_p = \sqrt{\vec{p}\cdot\vec{p}+m^2}$ and taking the cases $t'-t > 0$ and $t'-t < 0$ and the clockwise sense of the contour in the lower half plane we can evaluate the propagator by evaluating the contour integral according to Cauchy's integral formula [2].

$$2\pi i f(p_{0,1,2}) = \oint_{C_{1,2}} \frac{f(p_0)}{p_0 - p_{0,1,2}} dp_0$$

where on contour C_1 we have

$$f(p_0) = \frac{1}{(2\pi)} \frac{(\not{p}+m)e^{-ip_0(t'-t)+i\vec{p}\cdot(\vec{x}'-\vec{x})}}{(p_0 - E_p + i\epsilon)}$$

and on contour C_2 we have

$$f(p_0) = \frac{1}{(2\pi)} \frac{(\not{p} + m)e^{-ip_0(t'-t)+i\vec{p}\cdot(\vec{x}'-\vec{x})}}{(p_0 + E_p - i\epsilon)}$$

Note that the contour C_2 integration will incur an overall negative sign since it goes clockwise, but will be evaluated at $p_0 = E_P$, contour C_1 will incur a negative sign because it will pick the pole at $p_0 = -E_P$. Thus using Cauchy's integral formula and $p_0 = \pm E_p = \pm\sqrt{\vec{p}\cdot\vec{p} + m^2}$ appropriately to evaluate the contour integral in Eqn. 4.54. We do this and obtain

$$\begin{aligned} S_F^0(x'-x) &= -i\int \frac{d^3\vec{p}}{(2\pi^3)^3} \frac{e^{i\vec{p}\cdot(\vec{x}'-\vec{x})-iE_p(t'-t)}(E_p\gamma^0 - \vec{\gamma}\cdot\vec{p})}{2E_p}\theta(t'-t) \\ &+ i\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{e^{-i\vec{p}\cdot(\vec{x}'-\vec{x})+iE_p(t'-t)}(-E_p\gamma^0 - \vec{\gamma}\cdot\vec{p} + m)}{-2E_p}\theta(t-t') \end{aligned} \tag{4.56}$$

where the limit of $\epsilon \to 0$ has been taken. Note $\vec{p}$ in the second integral on the right of Eqn. 4.56 can been replaced by $-\vec{p}$ because the integral is over *all* $d^3\vec{p}$ and so is symmetric under $\vec{p} \to -\vec{p}$. The numerators of the two integrands on the right hand side then become $E_p\gamma^0 - \vec{\gamma}\cdot\vec{p} + m = \not{p} + m$ and $-E_p\gamma^0 + \vec{\gamma}\cdot\vec{p} + m = -\not{p} + m$ respectively. Recalling the positive and negative energy projection operators $\Lambda_\pm$ from Chapter 3 (Eqn. 3.135) we can rewrite the propagator in equation 4.56 in the following way

$$\begin{aligned} &S_F^0(x'-x) = \\ &-i\frac{m}{E_p}\int \frac{d^3\vec{p}}{(2\pi)^3}(\Lambda^+ e^{-ip\cdot(x'-x)}\theta(t'-t) + \Lambda^- e^{ip\cdot(x'-x)}\theta(t-t')) \end{aligned} \tag{4.57}$$

and can now clearly see that the propagator projects out positive energy solutions and moves them forward in time and projects out negative energy solutions and moves them backward in time-the latter are interpreted as positrons-this has been discussed at the very end of Chapter 3 and will be discussed again in this chapter. Using Eqns. 3.139 Eqns. 3.140 we can absorb the exponential and $\frac{1}{(2\pi)^3}$ factors in Eqn. 4.57 and write

$$\begin{aligned} S_F^0(x'-x) = -i\theta(t'-t)\int d^3\vec{p}\sum_{r=1}^{2}\psi_r^{E>0}(x')\bar{\psi}_r^{E>0}(x) \\ + i\theta(t-t')\int d^3\vec{p}\sum_{r=3}^{4}\psi_r^{E<0}(x')\bar{\psi}_r^{E<0}(x) \end{aligned} \tag{4.58}$$

where the plane waves

$$\begin{aligned}\psi^{E>0}_{r=1,2}(x) &= \sqrt{\frac{m}{E_p}}\frac{1}{(2\pi)^{3/2}}u^{E>0}_{s=1,2}(p)e^{-ip\cdot x}\\ \psi^{E<0}_{r=3,4}(x) &= \sqrt{\frac{m}{E_p}}\frac{1}{(2\pi)^{3/2}}v^{E<0}_{s=1,2}(p)e^{+ip\cdot x}\end{aligned} \tag{4.59}$$

have been defined as functions of x and x' in equation 4.58. The reader may ask about the normalization factors chosen in equation 4.59, recall that the normalization relations for just the spinors given in equations 3.121 involve a factor of $\frac{E}{m}$ (with $c = 1$) and so the factor of $\sqrt{\frac{m}{E}}$ here cancels that out ensuring that the integral of $\psi^\dagger\psi$ is the same over all space. The factor of $\frac{1}{(2\pi)^{\frac{3}{2}}}$ ensures delta function normalization for states of unequal momenta a characteristic that will be used soon in determining the action of this free propagator (Eqn. 4.57) on a wave function. Let us denote a wave function representing a positive energy solution of 4-momentum k upon which the propagator will act to be

$$\psi^{E_k>0}_{r=1,2}(x) = \sqrt{\frac{m}{E_k}}\frac{1}{(2\pi)^{3/2}}e^{-ik\cdot x}u^{E_k>0}_{1,2}(k) \tag{4.60}$$

Recall we had defined the retarded non-relativistic (Schrödinger) propagator to act in the following way in Eqn. 4.7

$$\theta(t'-t)\psi(x',t') = i\int d^3\vec{x}G_0(x',t';x,t)\psi(x,t) \tag{4.61}$$

$S^0_F(x'-x)$ has a somewhat different action, for the positive energy solution in Eqn. 4.60 we state (and verify immediately afterward) that

$$\begin{aligned}&\theta(t'-t)\psi^{E>0}_{r=1,2}(x',t') = i\int d^3\vec{x}S^0_F(x'-x)\gamma^0\psi^{E>0}_{r=1,2}(x) =\\ &i\int d^3\vec{x}S^0_F(x'-x)\gamma^0\sqrt{\frac{m}{E_k}}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-ik\cdot x}u^{E_k>0}_{1,2}(k)\end{aligned} \tag{4.62}$$

to verify Eqn. 4.62 we begin by using the propagator definition in Eqn. 4.57, and consider just the action of the first bit of the propagator on the plane wave (the factor of $i\times -i = 1$ has already been accounted for):

$$\begin{aligned}&\theta(t'-t)\iint d^3\vec{x}\frac{d^3\vec{p}}{(2\pi)^3}\frac{m}{E_p}\sqrt{\frac{m}{E_k}}\frac{e^{-i(p\cdot x')}e^{-i(\vec{k}-\vec{p})\cdot\vec{x}}}{(2\pi)^{\frac{3}{2}}}\\ &\times\Lambda_+\gamma^0 u^{E_k>0}_{s=1,2}(k)\end{aligned} \tag{4.63}$$

identifying $\int \frac{d^3\vec{x}}{(2\pi)^3} e^{-i(\vec{k}-\vec{p})\cdot\vec{x}}$ as the Dirac delta function we can write Eqn. 4.63 as

$$\begin{aligned}&\theta(t'-t)\int d^3\vec{p}\frac{m}{E_p}\sqrt{\frac{m}{E_k}}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{i(p\cdot x')}\delta^3(\vec{p}-\vec{k})e^{i(k^0-p^0)x^0}\\&\times\Lambda_+\gamma^0u^{E_k>0}_{s=1,2}(k)\end{aligned} \tag{4.64}$$

the integration over $d^3\vec{p}\delta^3(\vec{k}-\vec{p})$ sets

$$E_p = \sqrt{m^2+\vec{p}\cdot\vec{p}} = \sqrt{m^2+\vec{k}\cdot\vec{k}} = E_k$$

and we obtain

$$= \theta(t'-t)(\frac{m}{E_p})^{\frac{3}{2}}e^{-ip\cdot x'}\Lambda_+\gamma^0u^{E_p>0}_{s=1,2}(p) \tag{4.65}$$

using Eqn. 3.136 we know that $\Lambda_+\gamma^0u^{E_p>0}_{s=1,2}(p) = \frac{E_p}{m}u^{E>0}_{s=1,2}(p)$ and so we obtain after simplification:

$$\theta(t'-t)\sqrt{\frac{m}{E_p}}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-ip\cdot x'}u^{E_p>0}_{s=1,2}(p) = \theta(t'-t)\psi^{E_p>0}_{r=1,2}(x)$$

this verifies Eqn. 4.62. It should be clear that the term in $S^0_F(x'-x)$ with the Λ_- would have given 0 when operating on $\psi^{E_k>0}_{r=1,2}(x)$-the reader is encouraged to verify this. The same calculation for case of a negative energy solution $\psi^{E_k<0}_{r=3,4}(x)$ is left for the reader to verify. We can now summarize the action of the propagator for both negative and positive energy solutions:

$$\begin{aligned}i\int d^3\vec{x}S^0_F(x'-x)\gamma^0\psi^{E>0}_{r=1,2}(x) &= \theta(t'-t)\psi^{E>0}_{r=1,2}(x')\\ i\int d^3\vec{x}S^0_F(x'-x)\gamma^0\psi^{E<0}_{r=3,4}(x) &= -\theta(t-t')\psi^{E<0}_{r=3,4}(x')\end{aligned} \tag{4.66}$$

The reader should remember that one exponential in the propagator and in the wave function being acted combine with the integral over $\int d^4x$ yields a delta function in momentum which sets the momenta of the free particle wave to $\vec{p}$ after integration over $\int d^3\vec{p}$. After this the energy projection operator can act on the wave function and project out the appropriate state. Note that using the negative and positive energy solutions and representing the Fourier(equation 4.47) representation of the propagator as a contour integral with a *particular choice of countour* propagates an $E>0$ free wave as $t'\to\infty$, and an $E<0$ free wave as $t'\to-\infty$. Thus a negative energy particle is scattered from the future into the past, a positive energy solution scatters from the past into the future. We'll shortly relate the backward in time propagation of negative energy electrons to positrons. Recall section 3.5 where it was shown using discrete symmetries that negative energy solutions moving backward in time are equivalent to positive energy solutions of opposite charge moving forward in time, once the electromagnetic field is introduced in the next section this idea will be presented differently.

4.7 Computation of the Feynman propagator and S matrix in the presence of an electro-magnetic field.

Now that we have defined the properties of the free Feynman propagator we can proceed to introduce the electromagnetic potential $e\ \not{A}(x) = \gamma^\mu A_\mu(x) \times e$ as a perturbation, in a treatment similar to that for the non-relativistic propagator in the earlier sections of this chapter. The Feynman propagator in the presence of an electro-magnetic potential will be shown to describe processes in the relativistic regime. This means that the mass energy equivalence relation underlying the creation of observed positron and electron pairs, and their annihilation into photons must be described in a consistent manner by the formalism resulting from this propagator. Note that we had touched on the consequences of the negative energy solutions in section 3.10 of of chapter 3 and also identified the charge conjugated wave functions that describe positrons, we will use all of these concepts in this section and also in the final section where the relativistic S matrix will be derived. Onward with the formalism we remind the reader of Eqn. 4.46 where we had written down the following equation for the full propagator in the presence of an electro-magnetic 4-vector potential. This is reproduced below:

$$(i\,\not{\partial}' - e\,\not{A}(x') - m)S_F(x'-x) = \delta^4(x'-x)$$

which we rewrite as:

$$(i\,\not{\partial}' - m)S_F(x'-x) = \delta^4(x'-x) + e\,\not{A}(x')S_F(x'-x)$$

exploiting the properties of the Dirac δ function we can modify this further using an arbitrary dummy variable x_n denoting 3 space and 1 time co-ordinate:

$$(i\,\partial\!\!\!/' - m)S_F(x'-x) = \int d^4x_n \delta^4(x'-x_n)(\delta^4(x_n-x) + e\,A\!\!\!/(x_n)S_F(x_n-x)) \tag{4.67}$$

We can now write $\delta^4(x'-x_n)$ as $(i\,\partial\!\!\!/' - m)S_F^0(x'-x_n)$ and obtain

$$(i\,\partial\!\!\!/' - m)S_F(x'-x) = (i\,\partial\!\!\!/' - m)\int d^4x_n S_F^0(x'-x_n)(\delta^4(x_n-x) + e\,A\!\!\!/(x_n)S_F(x_n-x)) \tag{4.68}$$

as in the discussion preceding equation Eqn. 4.36 we note that if $A(x_n) = 0$ then $S_F(x'-x) = S_F^0(x'-x)$, bearing this in mind we note that we can remove $(i\,\partial\!\!\!/' - m)$ from both sides of Eqn. 4.68. Doing this and integrating $\delta^4(x'-x_n)S_F^0(x'-x_n)$ over d^4x_n we obtain the Lippman-Schwinger equation for the relativistic propagator (Feynman propagator):

$$S_F(x'-x) = S_F^0(x'-x) + e\int d^4x_n S_F^0(x'-x_n)\,A\!\!\!/(x_n)S_F(x_n-x) \tag{4.69}$$

The form of equation 4.69 is the same as that of the Lippman-Schwinger equation first shown in Eqn. 4.36. We can easily use this to generate a series solution for S_F in terms of $S_F^0(x'-x)$ and the potential and coupling. The reader is encouraged to carry out the procedure used for the non-relativistic case (Eqns. 4.36- 4.38) and it is easy to show that the series solution will be

$$S_F(x'-x) = S_F^0(x'-x) + e\int d^4x_1 S_F^0(x'-x_1)\,A\!\!\!/(x_1)S_F^0(x_1-x) \tag{4.70}$$

$$+e^2\iint d^4x_1 d^4x_2 S_F^0(x'-x_2)\,A\!\!\!/(x_2)S_F^0(x_2-x_1)\,A\!\!\!/(x_1)S_F^0(x_1-x)$$

$$+e^3\iiint d^4x_1 d^4x_2 d^4x_3 S_F^0(x'-x_3)\,A\!\!\!/(x_3)\cdots A\!\!\!/(x_1)S_F^0(x_1-x) + \cdots$$

to the desired order of accuracy. We note that we can define an exact wave function which is the solution to the Dirac equation in the presence of an electromagnetic 4-vector potential. We call this $\Psi(x')$ and note that it satisfies by its definition:

$$(i\,\partial\!\!\!/' - e\,A\!\!\!/' - m)\Psi(x') = 0 \tag{4.71}$$

by our own formalism this must involve the exact propagator:

$$\Psi(x') = i \int d^3\vec{x} S_F(x'-x)\gamma^0\psi(x) \tag{4.72}$$

where $\psi(x)$ is an initial wave. Using equation 4.69 and absorbing $i\gamma^0$, $S_F(x_n - x)$ by integrating over $d^3\vec{x}$ it is easy to see

$$\Psi(x') = \psi(x') + e\int d^4x_n S_F^0(x'-x_n) \not{A}(x_n)\Psi(x_n) \tag{4.73}$$

where $\psi(x')$ to the immediate right of the $=$ sign is a free wave propagated to x' and satisfies the homogenous equation $(i \not{\partial}' - m)\psi(x') = 0$. The second term in equation (4.73) represents the scattered wave and we examine this in the limits $t' \to \infty$ and $t' \to -\infty$. We note that $S_F^0(x'-x_n)$ describes free propagation from the n^{th} scattering to x'. We now look at the free propagator in equation (4.73), recall it is :

$$\begin{aligned} S_F^0(x'-x_n) = &- i\theta(t'-t_n)\int d^3\vec{p}\sum_{r=1}^{2}\psi_r^{E>0}(x')\bar{\psi}_r^{E>0}(x_n) + \\ & i\theta(t_n - t')\int d^3\vec{p}\sum_{r=3}^{4}\psi_r^{E<0}(x')\bar{\psi}_r^{E<0}(x_n) \end{aligned} \tag{4.74}$$

Note that in the limit $t' \to \infty$ only the positive energy (first) term of the propagator contributes and as $t' \to -\infty$ only the negative energy (second) term contributes-this is easily seen by considering the θ functions. If $t' \to \infty$ only the first piece contributes due to the $\theta(t'-t_n)$, note that $\not{A}(x_n)\Psi(x_n)$ in equation 4.73 is a 4-component object (CHECK WORDING) in which $A^\mu(x)$ can be represented by a Fourier transform as well (we shall see several examples of this in later chapters). The presence of the exponents from the wave function and the Fourier representation of the vector potential gives us a four component object and a delta function, exactly as in the derivation culminating in equation 4.66 the equivalizing of the momenta due to the delta function means that Λ_+ can then act on it, the resulting wave is then freely propagated to x'. In the limit $t' \to \infty$ only positive energy states propagate to the future and by analogy we can see that if $t' \to -\infty$ only negative energy states propagate to the past. Now we move on to calculating the S matrix containing the scattering amplitudes. In the limit $t' \to \infty$ a final *free* wave with $E > 0$ emerges whose spin and momentum we denote by the subscript $_f$:.

$$\psi_f(x') = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{m}{E_f}} u_{s_f=1,2}(p_f) e^{-ip_f \cdot x'} \tag{4.75}$$

A word of caution is neccesary here: we have taken the limit $t' \to \infty$ and we should remember that the emerging final state will be an $E > 0$ state, the label f that labels the momentum and spin denotes the state resulting from the propagation. Had we taken the limit $t' \to \infty$ a final $E < 0$ state would have emerged containing the effects of propagation backward in time-we'll consider this later.

To calculate the S-matrix we want to know the "overlap" between $\psi_f(x')$ and $\Psi(x')$. Note that $\Psi(x')$ is the wave that contains the effect of scattering and in the limit $t \to -\infty$ is the initial incoming *free* wave $\psi_i(x)$ represented by plane wave analogous to Eqn. (4.75).

$$\psi_i(x') = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{m}{E_i}} u_{s_i=1,2}(p_i) e^{-ip_i \cdot x'} \tag{4.76}$$

Note that the subscript i denotes the initial state and pertains to the momenta and spin. We want to calculate

$$S_{fi} = \lim_{t' \to \infty} \langle \psi_f(x') | \Psi(x') \rangle = \lim_{t' \to \infty} \int d^3\vec{x}' \psi_f^\dagger(x') \Psi(x') \tag{4.77}$$

Note that the wave $\Psi(x')$ is the total wave arriving at the space time point (x', t'), and it contains all the scattering that has taken place. Thus at time $t = -\infty$ we have a free wave which after undergoing scattering emerges as $\Psi(x')$ as $t' \to \infty$. Equation 4.77 then tells us the amplitude that it will be found in a free wave state $\psi_f(x)$. Inserting the incoming free wave $\psi_i(x)$ in place of $\psi(x)$ in equation 4.72 we then use equations and 4.73 to write:

$$\begin{aligned} S_{fi} &= \int d^3\vec{x}' \psi_f^\dagger(x') \psi_i(x') \\ &+ e \iint d^3x' d^4x_n \psi_f^\dagger(x') S_F^0(x' - x) \not{A}(x_n) \Psi(x_n) \end{aligned} \tag{4.78}$$

We note that the first term is simply a delta function in the difference in three momentum and the spinor indices of $\psi_f^\dagger$ and ψ_i. We obtain this result using the orthogonality relations in Eqn. 3.121 the plane wave factors and the definition of the Dirac delta function. The $\frac{e^{-ip_i \cdot x'}}{(2\pi)^{3/2}}$ and $\frac{e^{+ip_f \cdot x'}}{(2\pi)^{3/2}}$ functions combine with the integral over $d^3\vec{x}'$ to give a 3-dimensional δ function over 3-momentum using the orthogonality relations in Eqn. 3.121 we get $\frac{E}{m}$ times a Kronecker delta due to the spinor orthogonality relations. We a product of the Kronecker and $\delta^3(\vec{p}_f - \vec{p}_i) \times \delta_{fi}$ which we denote by δ_{fi} and we now have

$$S_{fi} = \delta_{fi} + e \iint d^3\vec{x}' d^4x_n \psi_f^\dagger(x') S_F^0(x' - x_n) \not{A}(x_n)\Psi(x_n) \tag{4.79}$$

where x_n is an arbitrary dummy variable. We'll now examine Eqn. 4.79 in the limits $t' \to \infty$ and after that $t' \to -\infty$. We now expand out $e \int d^3\vec{x}' \int d^4x_n \psi_f^\dagger(x') S_F^0(x' - x_n)$ only, explicitly keeping only the first term of $S_F^0(x' - x_n)$, suppressing $\theta(t' - t)$ which is now 1 since $t' \to +\infty$

$$\begin{aligned} &- ie \iint d^3\vec{x}' d^4x_n d^3\vec{p} \sqrt{\frac{m}{E_f}} \sqrt{\frac{m}{E_p}} \sum_{s_f,s}^{2} \frac{u_{s_f}^\dagger}{(2\pi)^{\frac{3}{2}}}(p_f) \frac{u_s(p)}{(2\pi)^{\frac{3}{2}}} e^{-i(p-p_f)\cdot x'} \\ &\times \bar{\psi}_{E>0}(x_n) \end{aligned} \tag{4.80}$$

the $\bar{\psi}(x_n)$ in the propagator (see equation 4.74) has been kept as it is and the $\psi(x')$ in the propagator is simply $\sqrt{\frac{m}{E_P}} \frac{1}{(2\pi)^{\frac{3}{2}}} u_s(p) e^{-ip\cdot x'}$. We can now spot our old friend the Dirac delta function in the integral $\int \frac{d^3\vec{x'}}{(2\pi)^3} e^{-i\vec{x'}\cdot(\vec{p}-\vec{p}_f)}$ and we can now rewrite equation 4.80

$$-ie \iint d^4x_n d^3\vec{p} \sqrt{\frac{m^2}{E_f E_p}} \sum_{s_f,s=1}^{2} u^\dagger(p_f) u_s(p) \delta^3(\vec{p} - \vec{p}_f) e^{-i(E-E_f)\cdot x'} \bar{\psi}_{E>0}(x_n)$$

and then integrating over $d^3\vec{p}$ setting $\vec{p} = \vec{p_f}$ we use the orthogonality relations of the spinors (equation 3.121)-we note that in the limit $t' \to +\infty$ Ψ_f would have to be a positive energy solution-we now obtain:

$$= -ie \int d^4x_n \left(\frac{m^2}{E_p^2}\right)^{1/2} \left(\frac{E_p}{m}\right) \bar{\psi}_{f,E>0}(x_n) = -ie \int d^4x_n \bar{\psi}_f(x_n)$$

and spin indices and momentum of $\bar{\psi}(x_n)$ consequently change to s_f, p_f. We now have the intermediate result:

$$S_{fi} = \delta_{fi} - ie \int d^4x_n \bar{\psi}_f(x_n) \not{A}(x_n) \Psi(x_n) \tag{4.81}$$

By considering the limit $t' \to -\infty$ we can easily obtain

$$S_{fi} = \delta_{fi} + ie \int d^4x_n \bar{\psi}_f(x_n) \not{A}(x_n)\Psi(x_n) \tag{4.82}$$

which would describe propagation backward in time to a negative energy state-the label $_f$ in equation 4.82 denotes it as a final state-the result of propagation-note however that as observers with a normal sense of time we will view this as an *initial* e^+ *state* so we have to choose our labels well.

We now rewrite down equation (4.73) using the subscript i to describe the initial state spin and momentum for the initial free state from which the scattered wave originates and propagates after undergoing a series of scatterings to space time point x_n

$$\Psi(x_n) = \psi_i(x_n) + e \int d^4x_n S_F^0(x' - x_n) \not{A}(x_n)\Psi(x_n) \tag{4.83}$$

where we have used the complete wave that has scattered upto x_{n-1}. We now proceed to iterate Eqn. 4.83 using this time another spacetime point x_{n-1}, followed by x_{n-2} , first substituting for $\Psi(x_n)$ in terms of $\Psi(x_{n-1})$ and then for $\Psi(x_{n-1})$ in terms of $\Psi(x_{n-2})$ and so on we obtain:

$$\begin{aligned} \Psi(x_n) &= \psi_i(x_n) + e \int d^4x_{n-1} S_F^0(x_n - x_{n-1}) \not{A}(x_{n-1})(\psi_i(x_{n-1}) \\ &+ e \iint d^4x_{n-1} d^4x_{n-2} S_F^0(x_n - x_{n-1}) \not{A}(x_{n-1}) S_F^0(x_{n-1} - x_{n-2}) \\ &\times \not{A}(x_{n-2})(\psi_i(x_{n-2}) + \cdots \end{aligned} \tag{4.84}$$

Substituting this for $\Psi(x_n)$ in equation 4.81 and setting $n = 1$ in the first term, $n = 2$ in the second etc the following series emerges:

$$\begin{aligned} S_{fi} &= \delta_{fi} - ie \int d^4x_1 \bar{\psi}_f(x_1) \not{A}(x_1)\psi_i(x_1) \\ &- ie^2 \iint d^4x_2 d^4x_1 \bar{\psi}_f(x_1) \not{A}(x_1) S_F^0(x_1 - x_2) \not{A}(x_2)\psi_i(x_2) - \cdots \end{aligned} \tag{4.85}$$

This development can be carried out for the $E < 0$ solutions in the limit $t' \to -\infty$, thus switching on the term containing $\theta(t_n - t')$ and propagation backward in time

$$\begin{aligned} S_{fi} &= \delta_{fi} + ie \int d^4x_1 \bar{\psi}_i(x_1) \not{A}(x_1)\psi_f(x_1) \\ &+ ie^2 \iint d^4x_1 d^4x_2 \bar{\psi}_i(x_1) \not{A}(x_1) S_F^0(x_1 - x_2) \not{A}(x_2)\psi_f(x_2) + \cdots \end{aligned} \tag{4.86}$$

Note once again that we have retained the order of appearance of the $_f$ and $_i$ labels and will interpret these wave functions for each physical case that we consider later.

The n^{th} order correction to the S-matrix can now be written in general form with $\epsilon_f = 1$ for *propagation to positive energy states in the far future* $t' \to +\infty$ and $\epsilon_f = -1$ for *propagation to negative energy states in the remote past or* $t' \to -\infty$

$$
\begin{aligned}
S_{fi} = -i\epsilon_f e^n \iint \cdots \int d^4x_n d^4x_{n-1} \cdots d^4x_1 \bar{\psi}_f(x_n) \not{A}(x_n) \\
\times S_F^0(x_n - x_{n-1}) \not{A}(x_{n-1}) S_F^0(x_{n-1} - x_{n-2}) \not{A}(x_{n-2}) \cdots \not{A}(x_1)\psi_i(x_1)
\end{aligned}
\tag{4.87}
$$

The potential $A^\mu(x)$ can for example be due to an external electric field due to a nucleus a proton, another electron, a positron, or free photons, the particle itself in the case of self energy. We will derive the results for cross sections for a wide variety of processes using the appropriate vector potential in equation 4.87. As a final note to this formalism we digress a little on terms with pieces such as $\cdots \int d^4x_{n-1} S_F^0(x_n - x_{n-1}) \not{A}(x_{n-2})\Psi(x_{n-1})$. As the reader will see in later chapters the vector potential $A_\mu(x_{n-1})$ can be expressed as a Fourier integral $\int d_4q e^{-iq\cdot x_{n-1}} A_\mu(q)$ since q has to have the units of inverse length it is easily identified as a 4 momentum, in this case we can think of $A_\mu(q)$ as the amplitude for the transfer of a particular momentum in the scattering processes. The reader should now recall the structure of the propagator and see that the plane wave exponential terms from $\Psi(x_{n-1})$ and the pieces of the propagator will combine to form a 4-dimensional Dirac delta function, this will set the propagator momentum p equal to $q + k$ where k is the 4-momentum of the initial wave $\Psi(x)$, the action of the projection operator(s) in S_F^0 will now act upon $(u_i(k)$ (the spinor part of $\Psi(x_{n-1}))$ and this will result in a superposition of waves with plane wave factor $e^{\pm i(p+q).x_n}$ propagating away from the space time point x_{n-1} thus the whole iterative series can be seen to result in the superposition of waves that have undergone several successive scatterings.

4.8 Interpreting the propagator series

The reader is asked to recall that the positron was discovered in 1933 validating the interpretation of the negative energy solutions of the Dirac equation by Dirac. In chapter 1 section 1.14 we showed using discrete symmetries that anti-particles travelling backward in time are equivalent to particles travelling forward in time.

To look at this issue perhaps more intuitively we can imagine a simple non-relativistic, non-radiating example of a particle propagating in constant electric field oriented in say the positive x direction. Let us assume that an observer with a "normal" sense of time observes the trajectory from time t_1 to t_2 ($t_2 > t_1$, since time flows forward) observes a change in momentum from intial momentum k to final momentum p along the x direction. The observer will conclude that the sign of the charge is the sign of $+\frac{p-k}{t_2-t_1}$. If time flows backward this motion begins with initial momentum of $-p$ and a final mometum of $-k$ at times t_2 and t_1 respectively with $t_2 > t_1$. The calculated acceleration with the backward flow of time will be $\frac{-k+p}{t_1-t_2} = -\frac{p-k}{t_2-t_1}$ and that the charge of the particle has the opposite to that of an observer with a normal sense of time.

With the discovery of the positron the need for a description of the processes of the appearance and dissappearance of electron-positron pairs arose. Since we now know that backward in time propagation of a particle with one charge when seen by an observer with an opposite (forward flowing) sense of time appears as a particle off opposite charge propagating forward in time, we can tie this up with our represetation the propagator. The reader is recommended the beautiful expositions given by Richard Feynman in texts and in journal articles [3 13 14]. To describe the physically observed processes of the appearance and annihilation of pairs of e^+ and e^- we first write down the expression for the S matrix again:

$$\begin{aligned} S_{fi} = -i\epsilon_f e^n \iint \cdots \int d^4x_n d^4x_{n-1} \cdots d^4x_1 \bar{\psi}_f(x_n) \not{A}(x_n) \\ \times S_F^0(x_n - x_{n-1}) \not{A}(x_{n-1}) S_F^0(x_{n-1} - x_{n-2}) \not{A}(x_{n-2}) \cdots \not{A}(x_1)\psi_i(x_1) \end{aligned} \tag{4.88}$$

Note that the terms contain free propagators ($S_F^0(x_{i+1} - x_i)$) and terms $\not{A}(x_i)$ sandwiched between initial and final state wave functions $\bar{\psi}_f(x_i)$ and $\psi_i(x_1)$ for an i^{th} order term. Note that since the free propagators $S_F^0(x_i - x_{i-1})$ and the potential $\not{A}(x_i)$ contain gamma matrices the terms in the series in Eqn. 4.88 will contain non-zero amplitudes for transitions between orthogonal $\psi_i(x)$ and $\psi_f(x)$ states (no surprise here) and non-zero amplitudes for transitions from $E > 0$ to $E < 0$ states and vice versa.

It is easy to see that the series in Eqn. 4.88 contains scatterings by a potential at space time points followed by a propagation to another space time point where another scattering takes place, of course the propagation can take place forward or backward in time. Lets now think of an electron

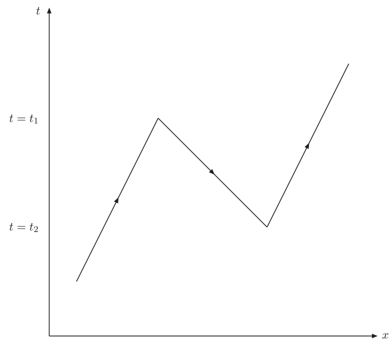

Fig. 4.3 The second order scattering of an electron in a potential forward and backward in time. The backward in time arrows are interpreted by an observer with a normal sense of time as positron moving forward in time.

propagating in a potential to second order as shown in Fig 4.3. We begin with an electron propagating forward in time, at time t_1 it is scattered backward in time to t_2 and then is again scattered forward in time. To an observer with a normal sense of time an electron appears to be propagating forward in time to t_1 and an oppositely charged positron arrives from t_2 at t_1 and at time t_1 we see the disappearance of an electron-positron pair. Conversely at time t_2 we see the emergence of an electron-positron pair with the electron propagating forward in time from t_2 onward. It's easy to see that the process conserves charge at each point in space time incorporating both positrons and electrons.

We'll now touch on how equation 4.88 can be used to describe scattering processes including the creation and annhilation of electron-positron pairs. The figures 4.4-4.7 contain simple graphical depictions of processes of scattering from one to state to the other including negative energy to positive energy states and vice-versa. The arrow shows the sequence in which the propagation is taken into account rather than the following the flow of time, the grey "blob" in the center of these figures represents all orders of scattering and this can include several succesive scatterings forward and backward in time before the emergence of a final state. These several scatterings represent the creation and annihilation of several virtual $e^+ - e^-$ (electron-positron) pairs. A description of these diagrams begins below:

(1) In figure 4.4 we see an electron with $E > 0$ propagating forward in time, this is a "vanilla" process, the electron is scat-

tered by a potential from an initial state of spin and 4-momentum p_i, s_i to p_f, s_f. The potential can be from a stationary or moving charged particle. To calculate the amplitude the appropriate potential, $\epsilon_f = +1$, and wave functions are inserted into equation 4.88, with $\psi_i(x_1) = \sqrt{\frac{m}{E_i.(2\pi^3)}} e^{-ip_i.x_1} u(p_i, s_i)$ carrying p_i, s_i and $\psi_f(x_n) = \sqrt{\frac{m}{E_f.(2\pi^3)}} e^{-ip_f.x_n} u(p_f, s_f)$. This represents the amplitude for an electron initially in a state with p_i, s_i to scatter into a state p_f, s_f.

(2) Lets now discuss the scattering of a positron scattering from a state of momentum and spin p_i, s_i to p_f, s_f. As we've noted this is equivalent to a negative energy electron going back in time, its "initial", incoming state has $-p_f, -s_f$ and "final", outgoing state has $-p_i, -s_i$. This is shown in figure 4.5. In equation 4.88 we replace $\bar{\psi}_f(x_n)$ with $\sqrt{\frac{m}{E_i.(2\pi^3)}} e^{-ip_i.x_n} \bar{v}(p_i, s_i)$ and $\psi_i(x_1)$ with $\sqrt{\frac{m}{E_f.(2\pi^3)}} e^{-ip_f.x_1} v(p_f, s_f)$ with $\epsilon_f = -1$.

(3) In figure 4.6 an $E > 0$ incoming electron propagating forward in time with p_-, s_- is scattered into an outgoing $E < 0$ state propagating (of course) backward in time with $-p_+, -s_+$. The backward in time motion of the negative energy electron is seen as a positron moving forward in time. The observer thus sees an $e^+ - e^-$ pair annihilate into photons represented by the potential A^μ. We have chosen the subscripts $_{+,-}$ to denote the charge of the particle seen by the observer. Following the arrows into and out of the potential we see that the amplitude can be calculated by setting $\epsilon_f = -1$, and replace $\psi_i(x_1)$ with $\sqrt{\frac{m}{E_-.(2\pi^3)}} e^{-ip_-.x_1} u(p_-, s_-)$ and $\bar{\psi}_f(x_n)$ with $\sqrt{\frac{m}{E_+.(2\pi^3)}} e^{-ip_+.x_n} \bar{v}(p_+, s_)$.

(4) In figure 4.7 an incoming $E < 0$ electron propagates backward in time and is scattered into an outgoing $E > 0$ state propagating forward in time. An observer with the usual sense of time sees a positive energy positron and positive energy electron propagating forward in time into their final states. This is the process of pair production. To calculate the appropriate amplitude in equation 4.88 we set $\epsilon_f = -1$, and replace $\psi_i(x_1)$ with $\sqrt{\frac{m}{E_+.(2\pi^3)}} e^{ip_+.x_n} v(p_+, s_+)$ and $\bar{\psi}_f(x_n)$ with $\sqrt{\frac{m}{E_-.(2\pi^3)}} e^{ip_-.x_1} \bar{u}(p_-, s_-)$.

(CHECK SPIN IN PREVIOUS DEVELOPMENT YOU'VE RE-

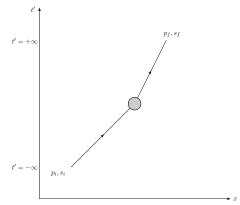

Fig. 4.4 An $E > 0$ electron is scattered by a potential from an initial state of momentum p_i, s_i to a final state of p_f, s_f.

VERSED ITS SIGN THROGHOUT IT SHOULD BE CONSISTENT WITH YOUR SPINOR DEFINITIONS)

The "blob" in these diagrams contains all orders of scattering. No process in this text is calculated beyond second order.

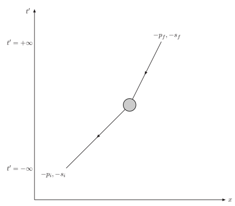

Fig. 4.5 An $E < 0$ electron scattering in a potential from an initial state of momentum $-p_f, -s_f$ to a final state of $-p_i, -s_i$, the propagation is backward in time, seen by an observer with the usual sense of time as a positron going from p_i, s_i to p_f, s_f

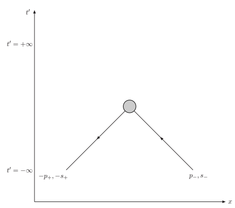

Fig. 4.6 An $E > 0$ electron with p_-, s_- is scattered by a potential into an $E < 0$ state of $-p_+, -s_+$ propagating backward in time. Seen by an observer with the usual sense of time as $e^+ - e^-$ pair annihilation.

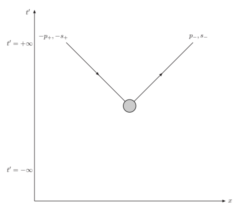

Fig. 4.7 An $E < 0$ electron with $-p_+, -s_+$ an propagates backward in time and is scattered by a potential and then propagates forward in time to p_i, s_i, seen by the observer as the creation of an $e^+ - e^-$ pair with p_+, s_+ and p_-, s_- respectively.

Chapter 5

Examples of Scattering Processes, Mott Scattering

In this chapter we will use the machinery we have developed and apply it to the problem of an electron scattering in the Coulomb field of a nucleus known as Mott scattering. This has been calculated in several other texts [9] [7] [12]. The steps in this calculation are conceptually similar to those used in other more involved calculations in the chapters to come. What follows is a very general description of the steps in the calculation. These are meant to introduce the reader to certain useful expressions and concepts that will be useful for their understanding of the remainder of this book as well as this chapter.

First we'll defining the incoming and outgoing particle free wave functions (appropriately normalized) that depend on the initial final state 4-momenta and spins respectively. A potential representing the nucleus will then be inserted into the series for the scattering amplitude developed at the end of Chapter 4 (Eqn. 4.88). The resulting amplitude will then be modulus squared which is a probability P. The entire process will be viewed as taking place in a large cubic box of volume V, side L and over a period of time T during which "everything" happens with the dimensions in space and time are assumed to be much larger than those within which the actual interaction between the electron and the nucleus (potential) itself takes place. We'll next weigh the probability for the possible final states of momentum that the scattered electron can transit to. In all calculations in this book we assume that we do not observe the initial spin state of the particles so these will be averaged and all final spin states that we observe will be summed. The time T and volume V that will help us interpret everything in terms of a rate $R = \frac{dP}{dt}$ or probability per unit time of scattering. It should be noted that the free particle incoming and outgoing wave functions will be normalized to 1 within volume V consistent with the definition of the duration and spatial extent of the scattering and its measurement. At the end of the day we will divide the rate by the incoming flux (probability per area per unit time) of particles to obtain a *differential cross-section* in units of area per solid exit angle. Given a flux of incoming particles defined as the modulus of the spatial part of the probability 4-current $\mid \vec{j}_i \mid$ (recall equation 2.29) we can write:

$$\sigma = \frac{R}{\mid \vec{j}_i \mid} \tag{5.1}$$

It should be clear from Eqn. 5.1 the cross section will have units of area. Cross sections are measured in barns with 1 barn = 10^{-24}cm^2. As mentionwe we will typically calculate the differential cross section per differential solid angle defined by the exit trajectory of a particle leaving the scattering region. As we shall see this solid angle will appear naturally in the counting of possible final states for the exiting particle to scatter into. It is hoped that all of this becomes intuitively clear to the reader in the following sections.

5.1 Mott Scattering: Setting up the scattering amplitude using the potential, initial and final wave functions

In the last chapter we obtained a series for the scattering amplitude of an electron in the presence of an external electromagnetic field which is reproduced below (chapter 4 equation 4.85

$$\begin{aligned}
S_{fi} = \delta_{fi} - ie \int d^4x_1 \bar{\psi}_f(x_1) \not{A}(x_1)\psi_i(x_1) \\
- ie^2 \int d^4x_1 d^4x_2 \bar{\psi}_f(x_2) \not{A}(x_2) S_F(x_2 - x_1) \not{A}(x_1)\psi_i(x_1) \\
- ie^3 \int d^4x_1 d^4x_2 d^4x_3 \bar{\psi}_f(x_3) \not{A}(x_3) S_F(x_3 - x_2) \not{A}(x_2) S_F(x_2 - x_1) \not{A}(x_1)\psi_i(x_1) \\
+ \cdots
\end{aligned} \tag{5.2}$$

please recall that this is for an electron, scattering forward in time and there is another series given in the same chapter for an electron scattering backward in time to represent a positron.

It should be clear that to calculate this series we require :

1. The initial and final state wave functions $\psi_i(x)$, $\psi_f(x)$
2. The four vector potential $A^\mu(x)$
3. If we were to go beyond first order (which we will not) we would need the free propagator $S_F^0(x_i - x_{i-1})$ defined in Chapter 4.

Let's first determine the free incoming and outgoing plane waves in Eqn. 5.2 representing the incident and scattered particles respectively. The initial and final plane waves using the subscripts i (initial) and f (final), both normalized to volume V (see opening remarks at the start of the chapter) are

$$\psi_i(x) = \sqrt{\frac{m}{E_i}}\frac{1}{\sqrt{V}}u(p_i, s_i)e^{-ip_i\cdot x} \tag{5.3}$$

$$\bar{\psi}_f(x) = \sqrt{\frac{m}{E_f}}\frac{1}{\sqrt{V}}\bar{u}(p_f, s_f)e^{ip_f\cdot x} \tag{5.4}$$

where the free particle spinor $u(p_i, s_i)$ and adjoint spinor $\bar{u}(p_i, s_i)$ were defined in chapter 3 (Eqns. 3.109).

This choice of normalization compensates for the side of the box when Lorentz contracted, thus the factor of $(\sqrt{\frac{m}{E_{i,f}}})^2$ (from $\psi^*\psi$) will provide the compensating factor of $\frac{1}{\gamma}$ to normalize the particle to 1 *within* the box of volume V in which "everything" takes place as mentioned in the opening remarks to this chapter. These wave functions will now go into the series in Eqn. 5.2. As stated before we will consider only the first order for this calculation. Recall (Eqn. 5.2) the first order term is simply

$$-ie\int d^4x\bar{\psi}_f(x)\not{A}(x)\psi_i(x) \tag{5.5}$$

where the dummy variable x has now taken the place of x_1 since there is no need to distinguish this from other terms in the series

What remains to be defined in Eqn. 5.5 is the four-vector potential. Since protons and neutrons ≈1860 times as heavy as the electron, we make the simplifying assumption that the nucleus doesn't move during the scattering. Thus if J^μ denotes the 4-vector current of the nucleus it has only the 0^{th} component and thus $\Box A^\mu(x) = 4\pi J^\mu(x)$ (see chapter 1 discussion preceding Eqn 1.76) has only a solution for the 0^{th} component of $A^\mu(x)$ which is a Coulombic potential.

Thus we can write the appropriate (0^{th}) component of the 4-vector potential :

$$A^\mu(x) = (\phi, 0, 0, 0) = \left(\frac{Ze}{|\vec{x}|}, 0, 0, 0\right) = A^0(x)$$

and

$$\not{A}(x) = \gamma_\mu A^\mu(x) = A_0(x)\gamma^0$$

so $\not{A}(x)$ is simply equal to $\gamma^0 Ze/|\vec{x}|$. The charge of the nucleus is $-Ze$ where e is the electron charge with $e < 0$.

using the explicit form of the 4-vector potential in the scattering amplitude and we obtain:

$$S_{fi} = -ie\sqrt{\frac{m^2}{E_i E_f V^2}} \int d^4x e^{ip_f \cdot x} \bar{u}(p_f, s_f)\gamma^0 \frac{Ze}{|\vec{x}|} e^{-ip_i \cdot x} u(p_i, s_i) \tag{5.6}$$

and collecting terms to be integrated over in x we rewrite this as:

$$S_{fi} = \frac{-ie^2 Zm}{\sqrt{E_i E_f} V} \bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i) \int \frac{d^4x e^{i(p_f - p_i)\cdot x}}{|\vec{x}|} \tag{5.7}$$

We can now break up the integral in S_{fi} into its space and time parts:

$$\int \frac{d^4x}{|\vec{x}|} e^{i(p_f - p_i)\cdot x} = \int dx_0 e^{i(E_f - E_i)x_0} \times \int d^3\vec{x} \frac{e^{-i(\vec{p}_f - \vec{p}_i)\cdot \vec{x}}}{|\vec{x}|}$$

The reader is reminded that one of the representations of the Dirac delta function is :

$$\delta(x - x') = \frac{1}{2\pi} \int dx e^{-ip\cdot(x - x')} \tag{5.8}$$

since p having inverse the units of x which is here a single variable. Generalized to n dimensions this is simply:

$$\delta^n(x - x') = \frac{1}{(2\pi)^n} \int d^n x e^{-i(x - x')\cdot p} \tag{5.9}$$

where the exponent now contains an n-dimensional scalar product–this n-dimensional representation is not needed for this process but we note it here for use in future calculations (both x and x' are 4-vectors). Since the Dirac Delta function is symmetric in its arguments we can identify

$$\int dx_0 e^{i(E_f - E_i)x_0} = 2\pi\delta(E_f - E_i)$$

, since energy has dimensions of inverse time. Using the delta function and denoting the change in 3-momentum of the scattering electron by $\vec{q} = \vec{p}_f - \vec{p}_i$ we can rewrite the above matrix element in Eqn. 5.7 as:

$$S_{fi} = \frac{-ie^2 Zm}{\sqrt{E_i E_f} V} \bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i) 2\pi\delta(E_f - E_i) \int d^3x \frac{e^{-i\vec{q}\cdot\vec{x}}}{|\vec{x}|} \quad (5.10)$$

Note the presence of the delta function in energy in Eqn. 5.10 means that at some point we have to integrate over dE_f to obtain a physically meaningful result. We now proceed by performing the integration over x in Eqn. 5.10 which is the Fourier transform of the Coulomb potential. This is solved by first writing it as an integral in spherical co-ordinates with $\vec{q} \cdot \vec{x} = \mid \vec{q} \mid r \cos\theta$:

$$\begin{aligned} &\int d^3x \frac{e^{-i\vec{q}\cdot\vec{x}}}{|\vec{x}|} \\ &= \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta) \int_0^{\infty} r^2 dr \frac{e^{-i|\vec{q}|r\cos\theta}}{r} \end{aligned} \quad (5.11)$$

cancelling one power of r and then introducing an exponential factor $e^{-\mu r}$ where we will take the limit $\mu \to 0$ at the end of our calculation we can write:

$$= \lim_{\mu\to 0} \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta) \int_0^{\infty} r dr e^{-i|\vec{q}|r\cos\theta - \mu r} \quad (5.12)$$

integrating over $d\phi$ and $d(\cos\theta)$ we obtain

$$\begin{aligned} &\lim_{\mu\to 0} -2\pi \int_0^{\infty} r dr \frac{(e^{-i|\vec{q}|r - \mu r} - e^{i|\vec{q}|r - \mu r})}{ir \mid \vec{q} \mid} \\ &= \lim_{\mu\to 0} \frac{-2\pi}{i \mid \vec{q} \mid} \times (\frac{1}{-i \mid \vec{q} \mid - \mu} - \frac{1}{i \mid \vec{q} \mid - \mu}) \\ &= \frac{-4\pi}{\mid \vec{q} \mid^2} \end{aligned} \quad (5.13)$$

and so

$$\int d^3x e^{i\vec{q}\cdot\vec{x}} \frac{Ze}{\mid \vec{x} \mid} = -\frac{4\pi Ze}{|\vec{q}|^2} \quad (5.14)$$

Thus the matrix element becomes:

$$S_{fi} = \frac{ie^2 Zm}{\sqrt{E_i E_f V}} \bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i) 2\pi\delta(E_f - E_i) \frac{(4\pi)}{|\vec{q}\,|^2} \qquad (5.15)$$

Now S_{fi} is a scattering amplitude for a given incoming and a given outgoing momentum and initial and final spins. This has to be modulus squared $| S_{fi} |^2$ and then this will represent a probability.

5.2 Mott Scattering: Counting all the possible final states, converting $| S_{fi} |^2$ into a probability

Note that the delta function in equation 5.15 imposes energy conservation and that given *one* initial $\vec{p}_i$ (initial momentum) there are *several* final momenta which satisfy $\delta(E_f - E_i)$, since the condition is only on the energy the final momentum can be oriented in any direction, as long as its magnitude helps satisfy the delta function via $E_f^2 = |\vec{p}_f|^2 + m^2$.

The probability represented by $| S_{fi} |^2$ must be modified to account the number of possible final momentum states available to the exiting electron. To count these we begin by representing the volume V as a large box of side L. Then all possible De-Broglie wavelengths of the electrons must be consistent with their being confined to this box. Note that since these are free particle states a *whole* number of wavelengths must fit into any one dimension of this box. The components of the momenta are related to the De-Broglie wavelengths by $p_{f,x} = \frac{h}{\lambda_x}$, $p_{f,y} = \frac{h}{\lambda_y}$ and $p_{f,z} = \frac{h}{\lambda_z}$ (we have reintroduced h so we can account for the factors of 2π associated with setting $\hbar = 1$). Relating this to the dimensions of the box: $n_x = \frac{L}{\lambda_x}$, $n_y = \frac{L}{\lambda_y}$ and $n_z = \frac{L}{\lambda_z}$. Here the $n_{x,y,z}$ tell us that a *whole number* of wavelengths must lie along each axis. We can simply write $n_x = \frac{p_{f,x}L}{h}$, $n_y = \frac{p_{f,y}L}{h}$, $n_z = \frac{p_{f,z}L}{h}$ or $n_x = \frac{p_{f,x}L}{2\pi\hbar}$, $n_y = \frac{p_{f,y}L}{2\pi\hbar}$, $n_z = \frac{p_{f,z}L}{2\pi\hbar}$ reverting to our choice of units where $\hbar = 1$ and so

$$n_x = \frac{p_{f,x}L}{2\pi}, \qquad n_y = \frac{p_{f,y}L}{2\pi}, \qquad n_z = \frac{p_{f,z}L}{2\pi}$$

.

Now let us think about what these ns mean. They are simply the number of wavelengths one can fit along a side of this box. If one changes the final momentum $\vec{p}_f$ by $\Delta\vec{p}_f$, this will correspond to changes in the number of wavelengths denoted by Δn_x, Δn_y and Δn_z each along the appropriate axis. Then for a variation in the value of p_x, p_y and p_z we have $\Delta n_x \Delta n_y \Delta n_z$ possibilities :

$$\Delta N = \Delta n_x \Delta n_y \Delta n_z = \frac{L^3}{(2\pi)^3} \Delta p_{f,x} \Delta p_{f,y} \Delta p_{f,z}$$

$$\therefore dN = d^3 n = \frac{L^3}{(2\pi)^3} d^3 \vec{p}_f$$

in terms of differentials, since $\vec{p}_f$ can be varied continuously. At this point we note that the volume of this box in which "everything happens" is $V = L^3$ and we can write:

$$dN = \frac{L^3}{(2\pi)^3} dp_{f,x} dp_{f,y} dp_{f,z} = \frac{V}{(2\pi)^3} d^3 \vec{p}_f$$

and so the probability for a particle to transit to a final state of momentum $\vec{p}_f$ is

$$dP = |S_{fi}|^2 \frac{V}{(2\pi)^3} d^3 \vec{p}_f \tag{5.16}$$

It is easy to see that for any scattering problem there will be a multiplicative factor of

$$\frac{V}{(2\pi)^3} d^3 \vec{p}_{f,k} \tag{5.17}$$

for each k^{th} *particle* exiting a scattering process. We have only one counting factor here since the nucleus is assumed to remain stationary in its final state.

Writing out the whole term with the explicit form of the amplitude multiplied by its complex conjugate:

$$P = \frac{Z^2 e^4}{V^2 |\vec{q}|^4} \frac{m_0^2}{E_i E_f} |\bar{u}(p_f, s_f) \gamma^0 u_i(p_i, s_i)|^2 (4\pi)^2 [2\pi \delta(E_f - E_i)]^2 \frac{V d^3 \vec{p}_f}{(2\pi)^3} \tag{5.18}$$

note that one power of the volume V will cancel and that the probability is proportional to the square of the delta function.

5.3 Mott Scattering: The square of the Dirac Delta function and the rate of transitions to the final state

We now deal with the fact that the probability for scattering contains the square of a Dirac Delta Function. Slightly simplifying 5.18 we obtain

$$P = \frac{Z^2 e^4}{V|\vec{q}|^4} \frac{m_0^2}{E_i E_f} |\bar{u}(p_f, s_f)\gamma^0 u_i(p_i, s_i)|^2 (4\pi)^2 [2\pi\delta(E_f - E_i)]^2 \frac{d^3\vec{p}_f}{(2\pi)^3} \quad (5.19)$$

where the reader is reminded that:

$$\delta(E_f - E_i) = \int_{-\infty}^{\infty} \frac{dt}{2\pi} e^{i(E_f - E_i)t} \quad (5.20)$$

We now begin to deal with the square of the Dirac Delta function in 5.19. It is important to remember that the Dirac delta function is not a true function, only its *integral* is defined. The same will be true for the square of the Dirac delta function. Keeping this in mind we begin to examine how the integral of the square of a Dirac-delta function might behave. Recall that by the definition of a Dirac-delta function for any arbitrary function $f(E_i)$

$$\int_{-\infty}^{\infty} \delta(E_f - E_i) f(E_i) dE_i = f(E_f)$$

We now consider an integral over the square of the Delta-function multiplied by $f(E_i)$:

$$\int_{-\infty}^{\infty} \delta(E_f - E_i)\delta(E_f - E_i) f(E_i) dE_i$$

Since the effect of multiplying a Dirac-delta function by another function and integrating it is to evaluate the function at the zero of the Dirac-delta function we can treat $\delta(E_f - E_i) f(E_i)$ together as one function and integrate over the remaining Dirac delta function to obtain :

$$\delta(E_f - E_f).f(E_f) = \delta(0) f(E_f)$$

as the result of the integration. In order to understand $\delta(0)$ we rewrite the Dirac-delta function 5.20 and integrate it changing the limits to $\pm T/2$ with the understanding that T tends to infinity and then take the limit of $E_f \to E_i$

$$\lim_{T\to\infty} \int_{-T/2}^{T/2} \frac{dt}{2\pi} e^{i(E_f - E_i)t} = \lim_{T\to\infty} \frac{1}{2\pi} \frac{1}{i(E_f - E_i)} \left\{ e^{i(E_f - E_i)T/2} - e^{-i(E_f - E_i)T/2} \right\}$$

$$= \lim_{T\to\infty} \left\{ \frac{1}{i(E_f - E_i)} \frac{1}{2\pi} 2i \sin(E_f - E_i)T/2 \right\} = \lim_{T\to\infty} \frac{1}{2\pi} \frac{2\sin(E_f - E_i)T/2}{(E_f - E_i)}$$

. Recall that a Dirac-delta function by definition is zero everywhere except where its argument goes to zero. Keeping this in mind we denote $E_f - E_i$ by ϵ and so the previous expression is simply

$$\frac{1}{\epsilon\pi}\sin(\frac{\epsilon T}{2})$$

where the limit $T \to \infty$ is understood. Taylor expanding the sine in a series we have:

$$\frac{1}{\pi\epsilon}(\frac{\epsilon T}{2} - \frac{\epsilon^3 T^3}{8.3!} + \cdots .)$$

It is clear that in the limit $E_f \to E_i$, $\epsilon \to 0$ and the only surviving term in the series is the first which means that

$$\lim_{E_i \to E_f} \delta(E_f - E_i) = \frac{T}{2\pi}$$

where it is understood that the limit $T \to \infty$ has to be taken.

Thus we simply replace $(2\pi\delta(E_f - E_i))^2$ by $(2\pi)^2\frac{T}{2\pi}\delta(E_f - E_i)$ with the understanding that the T will tend to infinity and that we have to promise to integrate over the remaining Dirac-delta function to obtain a physically meaningful result-at an appropriate point in the calculation. We note for future reference that a square of a Dirac-delta function in spatial momentum would be treated in an identical fashion starting from the representation (the exponent must be unitless, E has dimensions of $\frac{1}{T}$ and each component of $\vec{p}$ of $\frac{1}{L}$):

$$\delta(p_{f,x} - p_{i,x}) = \int_{-\infty}^{\infty} \frac{dx}{2\pi} e^{i(p_{f,x} - p_{i,x})\cdot x} = \lim_{L\to\infty} \int_{-L/2}^{L/2} \frac{dx}{2\pi} e^{i(p_{f,x} - p_{i,x})\cdot x} \quad (5.21)$$

It is clear that this will lead us to

$$(\delta(p_{f,x} - p_{i,x}))^2 = \frac{L}{2\pi}\delta(p_{f,x} - p_{i,x})$$

. This is in fact what we will obtain for the square of a Dirac-delta function in any spatial momentum.

As an aside we note that we will encounter these later in other scattering processes so at this point we write down an expression for the square of the four dimensional Dirac-delta function (forcing the conservation of four-momentum) for use in future chapters:

$$
\begin{aligned}
(\delta^4(p_f - p_i))^2 &= (\delta(E_f - E_i)\delta(p_{f,x} - p_{i,x})\delta(p_{f,y} - p_{i,y})\delta(p_{f,z} - p_{i,z}))^2 \\
&= \frac{T}{2\pi}\frac{L^3}{(2\pi)^3}\delta(E_f - E_i)\delta(p_{f,x} - p_{i,x})\delta(p_{f,y} - p_{i,y})\delta(p_{f,z} - p_{i,z}) \\
&= \frac{T}{2\pi}\frac{L^3}{(2\pi)^3}\delta^4(p_f - p_i) \\
&= \frac{T}{2\pi}\frac{V}{(2\pi)^3}\delta^4(p_f - p_i)
\end{aligned}
\tag{5.22}
$$

After this interlude we return to Eqn. 5.19 and replacing $[2\pi\delta(E_f - E_i)]^2$ by $(2\pi)^2\frac{T}{2\pi}\delta(E_f - E_i)$ and write down:

$$
P = \frac{Z^2e^4}{V|\vec{q}|^4}\frac{m^2}{E_iE_f}|\bar{u}(p_f, s_f)\gamma^0 u_i(p_i, s_i)|^2(4\pi)^2(2\pi)^2\frac{T}{2\pi}\delta(E_f - E_i)\frac{d^3\vec{p}_f}{(2\pi)^3}
\tag{5.23}
$$

Simplifying Eqn. 5.23 somewhat we can write:

$$
P = 4\frac{Z^2e^4}{V|\vec{q}|^4}\frac{m^2}{E_i}|\bar{u}(p_f, s_f)\gamma^0 u_i(p_i, s_i)|^2\frac{d^3\vec{p}_f}{E_f}\delta(E_f - E_i)T
\tag{5.24}
$$

We still have the offending factor of T which is undefined. It seems that the probability per particle transiting to a particular final state of momentum grows with time. The right hand side of Eqn. 5.24 also contains a factor of volume V which we have to interpret in terms of a physical measurable as well. Dealing with T first: lets think about what our experimental setup is like. We have an incoming particle flux impingent on a target and the flux is the probability per unit area per time. We can divide both sides by T and find the rate of particles transiting to final state f:

$$
dR = \frac{4Z^2e^4m^2}{VE_i|\vec{q}|^4}|\bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i)|^2\frac{d^3\vec{p}_f}{E_f}\delta(E_f - E_i)
\tag{5.25}
$$

We do need to get rid of the remaining Dirac-delta function. We do this by first changing the integral over $d^3\vec{p}_f$ to one over dE_f. We know that

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

and differentiating this with respect to $d|\vec{p}_f|$ we obtain

$$
\begin{aligned}
&2E_f\frac{dE_f}{d|\vec{p}_f|} = 2|\vec{p}_f| \\
&\to E_f dE_f = |\vec{p_f}|d|\vec{p}_f|
\end{aligned}
\tag{5.26}
$$

We can write $d^3\vec{p}_f$ in terms of spherical co-ordinates and recognize that $d^3\vec{p}_f = |\vec{p}_f|^2 d|\vec{p}_f| d\Omega_f$ where $d\Omega_f$ is the differential solid angle in the direction of the exiting particle's momentum.

We now use

$$d^3\vec{p}_f = |\vec{p_f}|^2 d|\vec{p}_f| d\Omega_f \tag{5.27}$$

with

$$|\vec{p_f}| d|\vec{p}_f| = E_f dE_f \tag{5.28}$$

in 5.25 to obtain:

$$dR = \frac{4Z^2e^4m^2}{VE_i|\vec{q}|^4} |\bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i)|^2 |\vec{p}_f| \frac{E_f}{E_f} dE_f d\Omega_f \delta(E_f - E_i) \tag{5.29}$$

Now this is simply the rate of transitions, we can simply integrate this over the Dirac-function in initial and final energy, the effect of this integration is merely to set $E_f = E_i$ in the remainder of the expression. We simply have to ensure that in the remainder of the calculation that the initial and final energies are equal.

$$dR = \frac{4Z^2e^4m^2}{VE_f|\vec{q}|^4} |\bar{u}(p_f, s_f)\gamma^0 u(p_i, s_i)|^2 |\vec{p}_f| d\Omega_f \tag{5.30}$$

or by writing $u(p_f, s_f)$ as u_f etc:

$$dR = \frac{4Z^2e^4m^2}{VE_f|\vec{q}|^4} |\bar{u}_f \gamma^0 u_i|^2 |\vec{p}_f| d\Omega_f \tag{5.31}$$

Note that since $E_i = E_f = \sqrt{| \vec{p_i} |^2 + m^2} = \sqrt{| \vec{p_f} |^2 + m^2}$ and so $| \vec{p_f} | = | \vec{p_i} |$ we will later denote $| \vec{p_{i,f}} |$ by $| \vec{p} |$ as the reader will see later.

5.4 Mott Scattering: Converting the differential rate to a differential cross section using the incoming flux

We now remind ourselves of how scattering experiments are set up: we have a beam of incoming particles of known flux impingent on a target and a detector to measure the scattered flux some distance away from the target. To test the strength and nature of interactions we will assume in this book that we will measure the scattered flux into solid angle $d\Omega_f$ subtended at the target along the line of the scattered particle's momentum. This defines the differential cross section per differential solid angle. We first define simply the differential cross section in terms of the incident flux of particles $\mid \vec{j}_i \mid$. The differential cross section using Eqn. 5.1

$$d\sigma \times \mid \vec{j}_i \mid = dR$$

where dR is the number of transitions per unit time. Here dR is directly proportional to $d\Omega_f$, thus we have an expression for $\frac{d\sigma}{d\Omega_f}$ and this allows us to calculate the expected number of particles per segmented detector element:

$$N_{\Delta\Omega_f} = \int_{\Delta\Omega_f} \mid \vec{j}_i \mid \frac{d\sigma}{d\Omega_f} d\Omega_f$$

Now $\mid \vec{j}_i \mid$ represents the flux of incident particles. Since we're setting up the experiment we can do this in any way that we want. We assume that the flux of particles is along the z axis and so recalling the discussion preceding and including Eqn. 3.55 we write the flux for a single particle moving along the positive z direction:

$$\mid \vec{j}_i \mid = c\bar{\psi}_i \gamma^3 \psi_i$$

of course $c = 1$ in our units and γ^3 is sandwiched between the spinors. We can evaluate this explicitly

$$\bar{\psi}_i \gamma^3 \psi_i = \frac{m}{EV} \frac{E+m}{2m} \begin{pmatrix} 1 & 0 & -\frac{p_z}{E+m} & 0 \end{pmatrix} \left(\begin{array}{c|c} 0 & \sigma_3 \\ \hline -\sigma_3 & 0 \end{array} \right) \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ 0 \end{pmatrix} \tag{5.32}$$

where we have used the definition of spinor and adjoint spinor (Eqns. 3.109-3.123) with the x and y components of momentum set to 0 and have used the cancellation of the plane wave exponential factors.

By explicit mutiplication it is easy to see that Eqn. 5.32 will yield $\frac{p_z}{EV}$ thus using p_i to denote the spatial momentum of an incident particle we obtain:

$$\mid \vec{j}_i \mid = \frac{|\vec{p}_i|}{E_i V} = \frac{v_i}{V} \tag{5.33}$$

it is easy to generalize expression 5.33 for a flux where beams of particles are incident on each other denoting their identities by subscripts $_1$ and $_2$

we have as a starting point for the flux calculation in the next chapter on electron-electron scattering:

$$| \vec{j}_i |= \frac{1}{V} \mid \frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2} \mid = \frac{\mid \vec{v}_1 - \vec{v}_2 \mid}{V} \tag{5.34}$$

Back to Mott scattering, using the flux for a single incident particle in Eqn. 5.33 we obtain:

$$d\sigma = \frac{dR}{| \vec{j}_i |} = \frac{4Z^2e^4m^2}{VE_f|\vec{q}|^4}\left(\frac{E_f V}{|\vec{p}_i|}\right)|\bar{u}_f\gamma^0 u_i|^2|\vec{p}_f|d\Omega_f$$

once again $|\vec{p}_f| = |\vec{p}_i|$ using $E_f = E_i$ we obtain:

$$\frac{d\sigma}{d\Omega_f} = \frac{4Z^2e^4m^2}{|\vec{q}|^4}|\bar{u}_f\gamma^0 u_i|^2 \tag{5.35}$$

We now have to calculate what $|\vec{q}|^4$ and what $|\bar{u}_f\gamma^0 u_i|^2$ are.

The easier piece first: $\vec{q} = \vec{p}_f - \vec{p}_i$ where $|\vec{p}_f| = |\vec{p}_i|$, thus they differ only in *direction*. Let us just denote $|\vec{p}_f|$ and $|\vec{p}_i|$ by $|\vec{p}|$.

Let us now assume that the target nucleus lies on the origin and the beam strikes it travelling along the z-axis going in the $+z$ direction and scatters off. Then the initial momentum is $\vec{p}_i = (0, 0, p)$ and the final $\vec{p}_f = (0, p\sin\theta, p\cos\theta)$ since the magnitude of the momentum doesn't change. Now $\vec{q} = \vec{p}_f - \vec{p}_i = (0, p\sin\theta, p\cos\theta - p)$ thus

$$|\vec{q}|^2 = \vec{q}\cdot\vec{q} = p^2\sin^2\theta + p^2\cos^2\theta + p^2 - 2p^2\cos\theta = 2p^2 - 2p^2\cos\theta \tag{5.36}$$

We now use the double angle formula to write $\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$ and replacing $\cos\theta$ in Eqn. 5.36 we have $|\vec{q}|^2 = 4p^2\sin^2\frac{\theta}{2}$ and finally we can also write

$$|\vec{q}|^4 = 16p^4\sin^4\frac{\theta}{2} \tag{5.37}$$

Our expression for the differential cross-section now becomes:

$$\frac{d\sigma}{d\Omega_f} = \frac{4Z^2e^4m^2}{16p^4\sin^4\frac{\theta}{2}}|\bar{u}_f\gamma^0 u_i|^2 \tag{5.38}$$

In the next section we examine the piece with the spinors:

$$|\bar{u}_f\gamma^0 u_i|^2 = \bar{u}_f\gamma^0 u_i(\bar{u}_f\gamma^0 u_i)^* \tag{5.39}$$

5.5 Averaging over unobserved initial spins and summing over final spin states: Casimir's Trick and trace theorems and the final differential cross section

In this section we will evaluate $|\bar{u}_f\gamma^0 u_i|^2$ whilst making certain assumptions about the nature of the experiment that is being performed. We assume that we *do not* observe the spins of the incoming or outgoing electrons. This assumption means that $|\bar{u}_f\gamma^0 u_i|^2$ will be averaged over the spins of the incoming particles and summed over the spins of the outgoing particles. As we shall see this problem actually reduces to taking the traces of products of gamma matrices contracted with momenta and multiplied by mass–this may seem like an obscure comment but will become obvious as the reader continues. This technique of reducing sums over spins to such products is known as Casimir's trick after the Dutch physicist Hendrik Casimir who developed this technique. We now proceed with the calculation.

We first calculate $|\bar{u}_f\gamma^0 u_i|^2 = \bar{u}_f\gamma^0 u_i(\bar{u}_f\gamma^0 u_i)^*$ and begin by considering $(\bar{u}_f\gamma^0 u_i)^*$. The reader is reminded that the fact that we have a γ^0 sandwiched between two spinors arises from the 4-vector potential $A^\mu(x) = (\phi(x), \vec{A}(x))$ having only the 0^{th} component $(\phi(x))$ since we have a static Coulomb potential. For a 4-potential due to a *moving* charged particle $A^\mu(x)$ we would have all four components which we would calculate by solving

$$\Box A^\mu(x) = 4\pi J^\mu(x)$$

see (discussion preceding and including Eqn. 1.76). Thus we would expect to find not just γ^0 but all gamma matrices γ^μ sandwiched between spinors.

Since later chapters will deal with scattering scenarios involving more than one moving charged particle (and photons) we can expect to have terms with *all* $A^\mu(x)$ components and all hence all γ^μ sandwiched between spinors. We might as well calculate $(\bar{u}_f\gamma^\mu u_i)^*$ and then use the specific case of $\mu = 0$ for Mott scattering. So we begin by writing:

$$(\bar{u}_f\gamma^\mu u_i)^* = (\bar{u}_{f,\sigma}\gamma^\mu_{\sigma\delta}u_{i,\delta})^* = \bar{u}^*_{f,\sigma}\gamma_{\sigma\delta}{}^{\mu *}u^*_{i,\delta} \tag{5.40}$$

in words the $\sigma\delta^{th}$ element of the complex conjugate of the γ matrix is summed over the δ element of the complex conjugate of the initial state spinor leaving free index σ which is then summed over the σ^{th} element of the complex conjugate of the final state conjugate spinor, thus forming the scalar product. Note that there are no free *matrix* or *spinor* free indices (there *is* a 4-vector free index μ) in this expression and it can be considered a scalar when considering the indices of other gamma matrices and spinors in the expression).

Recall the definition of the Hermitian conjugate of a matrix which is its transpose followed by complex conjugation thus:

$$\gamma^{\mu *}_{\sigma\delta} = \gamma^{\mu\dagger}_{\delta\sigma}$$

which gives us:

$$(\bar{u}_{f,\sigma}\gamma_{\sigma\delta}{}^{\mu}u_{i,\delta})^* = \bar{u}^*_{f,\sigma}\gamma^{\mu\dagger}_{\delta\sigma}u^*_{i,\delta} \tag{5.41}$$

It is easy to verify that

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$$

Keeping in mind the choice of indices we may write:

$$\gamma^{\mu\dagger}_{\delta\sigma} = \gamma^0_{\delta\theta}\gamma^\mu_{\theta\epsilon}\gamma^0_{\epsilon\sigma}$$

and substituting for $\gamma^{\mu\dagger}_{\delta\sigma}$ in 5.40 we have:

$$(\bar{u}_{f,\sigma}\gamma^\mu_{\sigma\delta}u_{i,\delta})^* = \bar{u}^*_{f,\sigma}\gamma^0_{\delta\theta}\gamma^\mu_{\theta\epsilon}\gamma^0_{\epsilon\sigma}u^*_{i,\delta} \tag{5.42}$$

Note that by definition $\bar{u}_\alpha = u^*_\beta\gamma^0_{\beta\alpha}$, since we have the complex conjugate of an element of $\bar{u}$ we can write: $\bar{u}^*_\alpha = (u^*_\beta\gamma^0_{\beta\alpha})^* = u_\beta\gamma^0_{\beta\alpha}$

(note the order of indices implies multiplication on left of the matrix by the spinor). Keeping this in mind we re-write Eqn. 5.42:

$$(\bar{u}_{f,\sigma}\gamma_{\sigma\delta}{}^{\mu}u_{i,\delta})^* = u_{f,\tau}\gamma^0_{\tau\sigma}\gamma^0_{\delta\theta}\gamma^\mu_{\theta\epsilon}\gamma^0_{\epsilon\sigma}u^*_{i,\delta} \tag{5.43}$$

re-ordering the terms:

$$(\bar{u}_{f,\sigma}\gamma_{\sigma\delta}{}^{\mu}u_{i,\delta})^* = u^*_{i,\delta}\gamma^0_{\delta\theta}\gamma^\mu_{\theta\epsilon}\gamma^0_{\tau\sigma}\gamma^0_{\epsilon\sigma}u_{f,\tau} \tag{5.44}$$

Since γ^0 is equal to its transpose $\gamma^0_{\tau\sigma}\gamma^0_{\epsilon\sigma} = \gamma^0_{\tau\sigma}\gamma^0_{\sigma\epsilon} = \delta_{\epsilon\tau}$ and $u^*_{i,\delta}\gamma^0_{\delta\theta}$ is by definition $= \bar{u}_{\theta,i}$ we can write (replacing index ϵ with τ):

$$(\bar{u}_{f,\sigma}\gamma_{\sigma\delta}{}^{\mu}u_{i,\delta})^* = \bar{u}_{i,\theta}\gamma^\mu_{\theta\tau}u_{f,\tau} \tag{5.45}$$

where in the last term we suppress the index, thus $(\bar{u}_f\gamma^\mu u_i)^* = \bar{u}_i\gamma^\mu u_f$ and $(\bar{u}_f\gamma^0 u_i)^* = \bar{u}_i\gamma^0 u_f$ and finally:

$$|\bar{u}_f \gamma^\mu u_i|^2 = \bar{u}_f \gamma^\mu u_i \bar{u}_i \gamma^\mu u_f \tag{5.46}$$

Now we do the promised average over all initial spins and sum over all possible final orientations of spin. For a single particle this is $\frac{1}{2}\times$(Sum over all possible final and initial spins). We now explicitly introduce all the indices in equation 5.39 and revert from using merely the subscripts f and i to explicitly using the momenta and spin arguments of the spinors.

$$\frac{1}{2} \sum_{\alpha,\beta,\delta,\sigma} \sum_{s_f, s_i} \bar{u}_\alpha(p_f, s_f) \gamma^0_{\alpha\beta} u_\beta(p_i, s_i) \bar{u}_\delta(p_i, s_i) \gamma^0_{\delta\sigma} u_\sigma(p_f, s_f)$$

$$= \frac{1}{2} \sum_{\alpha,\beta,\delta,\sigma} \sum_{s_f} \bar{u}_\alpha(p_f, s_f) \gamma^0_{\alpha\beta} \sum_{s_i} \underbrace{u_\beta(p_i, s_i) \bar{u}_\delta(p_i, s_i)} \gamma^0_{\delta\sigma} u_\sigma(p_f, s_f) \tag{5.47}$$

where the term above the braces is summed for each possible initial spin i. Thus s_i takes on values 1 and 2 and recall this is the $\beta\delta^{th}$ element of the matrix : $\left(\frac{\not{p}_i + m}{2m}\right)$ as we proved in Eqn. 3.138. From this point onward we will use these sums extensively in the calculation of spin averages and sums.

Since we have put in each index explicitly we can move things around:

$$\frac{1}{2} \sum_{\alpha,\beta,\delta,\sigma} \sum_{s_f} \bar{u}_\alpha(p_f, s_f) \gamma^0_{\alpha\beta} \left(\frac{\not{p}_i + m}{2m}\right)_{\beta\delta} \gamma^0_{\delta\sigma} u_\sigma(p_f, s_f)$$

and sum over β and δ

$$\frac{1}{2} \sum_{\alpha,\sigma} \sum_{s_f} \bar{u}_\alpha(p_f, s_f) \left\{\gamma^0 \frac{\not{p}_i + m}{2m} \gamma^0\right\}_{\alpha\sigma} u_\sigma(p_f, s_f)$$

Rearranging

$$\frac{1}{2} \sum_{\alpha,\sigma} \sum_{s_f} u_\sigma(p_f, s_f) \bar{u}_\alpha(p_f, s_f) \left\{\gamma^0 \frac{(\not{p}_i + m)}{2m} \gamma^0\right\}_{\alpha\sigma}$$

we now identify the sum over s_f to be the $\sigma\alpha^{th}$ element of $\left(\frac{\not{p}_f + m}{2m}\right)$– this is the sum over spins summarized in and in the discussion preceding Eqn. 3.139-3.140. We now have :

$$\frac{1}{2} \sum_{\alpha,\sigma} \left(\frac{\not{p}_f + m}{2m}\right)_{\sigma\alpha} \left(\gamma^0 \frac{\not{p}_i + m}{2m} \gamma^0\right)_{\alpha\sigma}$$

summing over α we get

$$= \frac{1}{2}\sum_{\sigma}\left\{\left[\frac{\not{p}_f + m}{2m}\right]\left[\frac{\gamma^0(\not{p}_i + m)\gamma^0}{2m}\right]\right\}_{\sigma\sigma}$$

the sum over σ of the $\sigma\sigma^{th}$ element of the matrix shown and is by definition the trace (denoted by Tr) :

$$= \frac{1}{2}Tr\left\{\left(\frac{\not{p}_f + m}{2m}\right)\left(\frac{\gamma^0(\not{p}_i + m)\gamma^0}{2m}\right)\right\}$$

We now have

$$\frac{d\sigma}{d\Omega_f} = \frac{1}{2}Tr\left[\left(\frac{\not{p}_f + m}{2m}\right)\left(\frac{\gamma^0(\not{p}_i + m)\gamma^0}{2m}\right)\right] \times \frac{4Z^2e^4m^2}{16p^4\sin^4(\frac{\theta}{2})}$$

Recall that the square of γ^0 is simply the identity, using this we need the trace of:

$$Tr(\frac{(\not{p}_f\gamma^0\ \not{p}_i\gamma^0 + m\ \not{p}_f + m\gamma^0\ \not{p}_i\gamma^0 + m^2)}{4m^2})$$

Where the m^2 term is understood multiplied by the identity $\mathbb{I} = (\gamma^0)^2$. Since the γ matrices are traceless the we know is that the trace of any one γ matrix is zero and so $m\ \not{p}_f$ contributes 0

Secondly we know that $Tr(\gamma^0\ \not{p}_i\gamma^0) = Tr(\gamma^0\gamma^0\ \not{p}_i)$ from the cyclic property of the trace, and since $(\gamma^0)^2$ is the identity,

$$Tr(\gamma^0\ \not{p}_i\gamma^0) = Tr(\gamma^0\gamma^0\ \not{p}_i) = Tr(\not{p}_i)$$

which is zero.

So we need only evaluate:

$$\frac{1}{4m^2}Tr(\not{p}_f\gamma^0\ \not{p}_i\gamma^0 + m^2) \tag{5.48}$$

Of these only the first one is actually difficult since the trace of m^2 times the identity is simply $4m^2$.

It should be obvious at this point that we need to make some statements about the traces of products of γ matrices. So let's first prove a few theorems, some of which will be used in later chapters:

Theorem I

The trace of a product of an odd *number of* γ *matrices is zero.*

Proof *We have already defined and used* $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ *in Chapter 3, Section 3.4 If the reader is brave enough to multiply it out it is* $\left(\begin{array}{c|c} 0 & \mathbb{I} \\ \hline \mathbb{I} & 0 \end{array}\right) = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right)$. *It is easy to verify that* $(\gamma^5)^2 = \mathbb{I}$.

Now note that

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

μ *has to be either* 0, 1, 2 *or* 3 *if we move the* γ^μ *past each of the* γ*s to its left it is clear from the anti-commutation relations that we will pick up a factor of -1 for every matrix it goes past* <u>*except when it encounters itself*</u> *which one of the 4 have to be. Hence we will have* 3 *factors of* (-1) *or 1 factor of* -1 *overall, therefore we can write:*

$$\gamma^5\gamma^\mu = -\gamma^5\gamma^\mu$$

this means we can simply write down the anticommutation relation for γ^5 *with anyone of the other* γ*s:*

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

After this we move on to the remainder of the proof: consider the product of an odd number of γ*s:* $\not{A}_1\ \not{A}_2\cdots\ \not{A}_n$ *this can always be written as* $\gamma^5\gamma^5\ \not{A}_1\ \not{A}_2\cdots\not{A}_n$ *since* $(\gamma^5)^2 = \mathbb{I}$

Taking the trace of such a product:

$$Tr(\not{A}_1\cdots\not{A}_n) = Tr(\gamma^5\gamma^5\ \not{A}_1\ \not{A}_2\cdots\not{A}_n)$$

where $\gamma^5\gamma^5 = \mathbb{I}$ *has been inserted. Using the cyclic property of the trace this equals*

$$Tr(\gamma^5\ \not{A}_1\ \not{A}_2\cdots\not{A}_n\gamma^5)$$

If the last γ^5 *is now moved to the left it will pick up a* $-$ *sign each time it moves past one of the gamma matrices* $\not{A}_n$*–in moving all the way to the left this will happen an odd number of times, this means*

$$Tr(\not{A}_1\cdots\not{A}_n) = Tr(\gamma^5\gamma^5\ \not{A}_1\ \not{A}_2\cdots\not{A}_n) = Tr(\gamma^5\ \not{A}_1\cdots\not{A}_n\gamma^5)$$

$$= (-1)^n Tr(\gamma^5\gamma^5\ \not{A}_1\cdots\not{A}_n) = -Tr(\gamma^5\gamma^5\ \not{A}_1\cdots\not{A}_n) = -Tr(\not{A}_1\cdots\not{A}_n)$$

Therefore

$$Tr(\not{A}_1\cdots\not{A}_n) = -Tr(\not{A}_1\cdots\not{A}_n) = 0$$

if n *is odd. QED.*

Theorem II

The trace of a product of $\not{a}\ \not{b}$ *is* $4a\cdot b$*.*

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

$$Tr(\not{a}\,\not{b}) = Tr(\gamma^\mu\gamma^\nu a_\mu b_\nu)$$

$$Tr[(g^{\mu\nu} - \gamma^\nu\gamma^\mu)a_\mu b_\nu] = Tr(2a\cdot b\mathbb{I}) - Tr\gamma^\nu\gamma^\mu a_\mu b_\nu$$

$$Tr(\not{a}\,\not{b}) = Tr(2a\cdot b\mathbb{I}) - Tr(\not{b}\,\not{a})$$

but $\mathbb{I}$ *is* 4×4

$$\therefore Tr(\not{a}\,\not{b}) + Tr(\not{b}\,\not{a}) = 8a\cdot b$$

but due to the cyclical property of trace $Tr(\not{a}\,\not{b}) = Tr(\not{b}\,\not{a})$.

$$\therefore 2Tr(\not{a}\,\not{b}) = 8a\cdot b \implies Tr(\not{a}\,\not{b}) = 4a\cdot b \ QED.$$

Theorem III

For a product of an even number of γ *matrices:*

$$Tr[\not{A}_1\,\not{A}_2\cdots\not{A}_n] = A_1\cdot A_2 Tr[\not{A}_3\cdots\not{A}_n] - A_1\cdot A_3 Tr[\not{A}_2\,\not{A}_4\cdots\not{A}_n]$$

$$+A_1\cdot A_4 Tr[\not{A}_2\,\not{A}_3\cdots\not{A}_n] - A_1\cdot A_5[\not{A}_2\,\not{A}_3\cdots\not{A}_n] + \ etc\cdots$$

We will prove this by moving $\not{A}_1$*s successively past the its neighbour to the right. This is done by using*

$$\not{A}_1\,\not{A}_i = \gamma^\mu\gamma^\nu A_{1\mu}A_{i\nu} = (2g^{\mu\nu} - \gamma^\nu\gamma^\mu)A_{1\mu}A_{i\nu} = 2A_1\cdot A_i - \not{A}_i\,\not{A}_1$$

where the order of appearance of the γ *matrix indices with respect and the space time indices of the* A_p *has been used. The proof below uses this succesively moving* $\not{A}_1$ *past each neighbour.*

Consider:

$$\not{A}_1\,\not{A}_2\cdots\not{A}_n = (2g^{\mu\nu} - \gamma^\nu\gamma^\mu)A_{1\mu}A_{2\nu}\,\not{A}_3\,\not{A}_4\cdots\not{A}_n$$

$$= 2A_1\cdot A_2(\not{A}_3\,\not{A}_4\cdots\not{A}_n) - \not{A}_2\underbrace{\not{A}_1\,\not{A}_3}\,\not{A}_4\cdots\not{A}_n$$

$$2A_1\cdot A_2(\not{A}_3\,\not{A}_4\cdots\not{A}_n) - \not{A}_2(2g^{\mu\nu} - \gamma^\nu\gamma^\mu)A_{1\mu}A_{3\nu}\,\not{A}_4\cdots\not{A}_n$$

$$2A_1\cdot A_2(\not{A}_3\,\not{A}_4\cdots\not{A}_n) - 2A_1\cdot A_3(\not{A}_2\cdots\not{A}_n) + \not{A}_2\,\not{A}_3\underbrace{\not{A}_1\,\not{A}_4}\cdots\not{A}_n$$

$$2A_1\cdot A_2(\not{A}_3\,\not{A}_4\cdots\not{A}_n) - 2A_1\cdot A_3(\not{A}_2\,\not{A}_4\cdots\not{A}_n) + \not{A}_2\,\not{A}_3(2g^{\mu\nu} - \gamma^\nu\gamma^\mu)A_{1\mu}A_{4\nu}\,\not{A}_5\cdots\not{A}_n$$

and ultimately after n *terms we'll have*

$$\not{A}_1\,\not{A}_2\cdots\not{A}_n = 2A_1\cdot A_2(\not{A}_3\,\not{A}_4\cdots\not{A}_n) - 2A_1\cdot A_3(\not{A}_2\,\not{A}_4\cdots\not{A}_n) + 2A_1\cdot A_4(\not{A}_2\,\not{A}_3\cdots\not{A}_n)$$

$-2A_1 \cdot A_5(\not{A}_2 \not{A}_3 \not{A}_4 \cdots \not{A}_n) + 2A_1 \cdot A_n(\not{A}_2 \cdots \not{A}_{n-1}) - \cdots \not{A}_2 \not{A}_3 \cdots \not{A}_n \not{A}_1$

now take the trace of both sides

$Tr(\not{A}_1 \not{A}_2 \cdots \not{A}_n) = 2A_1 \cdot A_2 Tr(\not{A}_3 \not{A}_4 \cdots \not{A}_n) - 2A_1 \cdot A_3 Tr(\not{A}_2 \not{A}_4 \cdots \not{A}_n) + \cdots - Tr(\not{A}_2 \not{A}_3 \cdots \not{A}_n \not{A}_1)$

but by the cyclic property of the trace the last term is $Tr(\not{A}_1 \not{A}_2 \cdots \not{A}_n)$ *and so*

$2Tr(\not{A}_1 \not{A}_2 \cdots \not{A}_n) = 2A_1 \cdot A_2 Tr(\not{A}_3 \not{A}_4 \cdots \not{A}_n) - 2A_1 \cdot A_3 Tr(\not{A}_2 \not{A}_4 \cdots \not{A}_n)$

and finally

$Tr(\not{A}_1 \not{A}_2 \cdots \not{A}_n) = A_1 \cdot A_2 Tr(\not{A}_3 \not{A}_4 \cdots \not{A}_n) - A_1 \cdot A_3 Tr(\not{A}_2 \not{A}_4 \cdots \not{A}_n) + \cdots$ *QED.*

Going back to the expression for differential cross-section:

$$\frac{d\sigma}{d\Omega_f} = \frac{4Z^2e^4m^2}{16p^4\sin^4(\frac{\theta}{2})} \times \frac{1}{2} Tr\left[\frac{1}{4m^2}(\not{p}_f\gamma^0 \not{p}_i\gamma^0 + m^2)\right]$$

we consider now $Tr(\not{p}_f\gamma^0 \not{p}_i\gamma^0)$ in light of Theorem III.

Let $b_\mu = (1,0,0,0)$ be a unit vector with an entry in the 0^{th} slot, then $Tr(\not{p}_f\gamma^0 \not{p}_i\gamma^0) = Tr(\not{p}_f \not{b} \not{p}_i \not{b})$

By Theorem III:

$Tr(\not{p}_f\gamma^0 \not{p}_i\gamma^0) = Tr(\not{p}_f \not{b} \not{p}_i \not{b}) = p_f \cdot b Tr(\not{p}_i \not{b}) - p_f \cdot p_i Tr(\not{b} \not{b}) + p_f \cdot b Tr(\not{b} \not{p}_i)$

Note by Theorem II $Tr(\not{a} \not{b}) = 4a \cdot b$

$$\implies Tr(\not{p}_f\gamma^0 \not{p}_i\gamma^0) = (p_f \cdot b)4(p_i \cdot b) - 4p_f \cdot p_i + p_f \cdot b4b \cdot p_i$$

$$4E_fE_i - 4(E_f \times E_i) + 4\vec{p}_f \cdot \vec{p}_i + 4(E_f \times E_i)$$

$$= 4E_fE_i + 4\vec{p}_f \cdot \vec{p}_i$$

$$4E_f^2 + 4\vec{p}_f \cdot \vec{p}_i = 4(E^2 + |\vec{p}|^2\cos\theta)$$

where we have used $Tr(m^2\mathbb{I}) = 4m^2$, $E_f = E_i = E$ and $|\vec{p_f}| = |\vec{p_i}| = |\vec{p}|$

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2e^4m^2}{16p^4\sin^4(\frac{\theta}{2})}\frac{1}{2}\left(\frac{4E_f^2}{4m^2} + \frac{4|\vec{p}|^2\cos\theta}{4m^2} + \frac{4m^2}{4m^2}\right) = \frac{4Z^2e^4}{16p^4\sin^4(\frac{\theta}{2})}\frac{1}{8}\left(4E_f^2 + 4|\vec{p}|^2\cos\theta + 4m^2\right)$$

Note that by using the double angle formula: $2\cos^2(\frac{\theta}{2}) - 1 = \cos\theta$ can be used to write $1 - 2\sin^2(\frac{\theta}{2}) = \cos\theta$

$$\frac{4Z^2e^4}{16p^4\sin^4(\frac{\theta}{2})} \times \frac{1}{8}\left(4E_f^2 + |\vec{p}|^2 4[1 - 2\sin^2(\frac{\theta}{2})] + 4m^2\right)$$

$$
\begin{aligned}
&\frac{4Z^2e^4}{16p^4\sin^4(\frac{\theta}{2})} \times \frac{1}{8}\left(4E_f^2 + 4|\vec{p}|^2 + 4m^2 - 8|\vec{p}|^2\sin^2(\frac{\theta}{2})\right) \\
&\frac{4Z^2e^4}{16p^4\sin^4(\frac{\theta}{2})} \times \frac{1}{8}[4E_f^2 + 4|\vec{p}|^2 + 4|E_f|^2 - 4|\vec{p}|^2 - 8|\vec{p}\sin^2(\frac{\theta}{2})] \\
&= \frac{4Z^2e^4}{16p^4\sin^4(\frac{\theta}{2})}[E_f^2 - |\vec{p}|^2\sin^2(\frac{\theta}{2})] \\
&= \frac{1}{4}\frac{Z^2e^4E_f^2}{p^4\sin^4(\frac{\theta}{2})}[1 - \frac{|\vec{p}|^2}{|E_f|^2}\sin^2(\frac{\theta}{2})] \\
&= \frac{1}{4}\frac{Z^2e^4}{\sin^4(\frac{\theta}{2})}\left(\frac{E_f^2}{p^4}\right)[1 - \beta^2\sin^2(\frac{\theta}{2})] \\
&= \frac{Z^2e^4}{4\sin^4(\frac{\theta}{2})\beta^2p^2}(1 - \beta^2\sin^2(\frac{\theta}{2}))
\end{aligned}
$$

This is called the Mott-scattering cross-section. Note the extra term proportional to $-\beta^2\sin^2(\frac{\theta}{2})$ which modifies what would simply have been the Rutherford scattering cross-section.

5.6 Coulomb scattering of positrons

We will now use the Mott scattering calculation to obtain a result for positron scattering in the Coulomb potential of a nucleaus. One way to do this is to recall that in our formalism we view positrons as negative energy electrons moving backward in time. Another would be to rewrite our formalism for scattering using the charge conjugate wave functions (see chapter 3) and doing everything in terms of positrons running forward in time. We use the former approach here and everywhere else in this manuscript.

We begin by rewriting the result for the iterative series for S_{fi} below for negative energy electrons which we first encountered in Chapter 4:

$$
\begin{aligned}
S_{fi} = \delta_{fi} &+ ie\int d^4x_1\bar{\psi}_f(x_1)\,\not{A}(x_1)\psi_i(x_1) \\
&+ ie^2\int d^4x_1d^4x_2\bar{\psi}_f(x_2)\,\not{A}(x_2)S_F(x_2 - x_1)\,\not{A}(x_1)\psi_i(x_1) \\
&+ ie^3\int d^4x_1d^4x_2d^4x_3\bar{\psi}_f(x_3)\,\not{A}(x_3)S_F(x_3 - x_2)\,\not{A}(x_2)S_F(x_2 - x_1)\,\not{A}(x_1)\psi_i(x_1) \\
&+ \cdots
\end{aligned}
\tag{5.49}
$$

The reader is reminded here that e remains the electron charge since we are viewing positrons as *electrons* running backward in time. The origin of the positive sign in (LEFTOVER CHECK THIS) the terms arises in fact from the contour integration in Chapter 4 which the reader is asked to revisit. Of course the charge conjugation operation can also be used (see the discussion in Chapter 3) to treat the problem as one of positrons travelling *forward* in time.

As stated however we use the picture of electrons with negative energy moving backward in time. We will only carry out this calculation to first order-just like the electron calculation. Note that the first step is to think of appropriate substitutions for $\psi_f(x_1)$ and $\psi_i(x_1)$ in Eqn. 5.49. For a scattering process that runs backward in time the label f denotes the free quantum state after the scattering has taken place. For a process that proceeds forward in time the propagator acts upon $\psi_{i,E>0}(x)$ and we take the scalar product of the resulting state with a free wave $\psi_{f,E>0}(x)$. We can see that for a scattering that proceeds backward in time

We have used the $E > 0$ and $E < 0$ labels to distinguish between the solutions in the above discussion. We drop these distinctions and state that:

1. In equation 5.49 we replace
$$\bar{\psi}_f(x_1) = \sqrt{\frac{m}{E_f}}\frac{1}{\sqrt{V}}\bar{u}(p_f, s_f)e^{ip_f\cdot x}$$
by
$$\sqrt{\frac{m}{E_i}}\frac{1}{\sqrt{V}}\bar{v}(p_i, s_i)e^{-ip_i\cdot x}$$
2. We also replace
$$\psi_i(x) = \sqrt{\frac{m}{E_i}}\frac{1}{\sqrt{V}}u(p_i, s_i)e^{-ip_i\cdot x}$$
by
$$\sqrt{\frac{m}{E_f}}\frac{1}{\sqrt{V}}v(p_f, s_f)e^{+p_f\cdot x}$$
(The reader is encouraged to refresh themselves by looking at the discussion preceding Eqn. 5.4)

Thus an incoming $E > 0$ electron has been replaced by an outgoing $E < 0$ electron (positron) and an outgoing $E > 0$ electron has been replaced by an incoming $E < 0$ electron (positron). We can see clearly that we can write down S_{fi} to first order trivially using Eqn. 5.15

$$S_{fi} = -\frac{ie^2 Zm}{\sqrt{E_i E_f}V}\bar{v}(p_i, s_i)\gamma^0 v(p_f, s_f)2\pi\delta(E_i - E_f)\int d^3x \frac{e^{-i\vec{q}\cdot\vec{x}}}{|\vec{x}|} \quad (5.50)$$

The squaring of the matrix elements and the counting of the final states, the division by the flux etc is all the same. The summing over final and averaging over initial spins however requires a slightly different step. We begin at the expression for the differential cross section which is trivial to get by appropriate modification of Eqn. 5.38

$$\frac{d\sigma}{d\Omega_f} = \frac{4Z^2 e^4 m^2}{16p^4 \sin^4 \frac{\theta}{2}}|\bar{v}_i \gamma^0 v_f|^2 \quad (5.51)$$

it is easy for the reader to check following the development between Eqns. 5.40 and Eqn. 5.47 will yield:

$$\frac{1}{2}\sum_{\alpha,\beta,\delta,\sigma}\sum_{s_f,s_i}\bar{v}_\alpha(p_i, s_i)\gamma^0_{\alpha\beta}v_\beta(p_f, s_f)\bar{v}_\delta(p_f, s_f)\gamma^0_{\delta\sigma}u_\sigma(p_i, s_i)$$

$$= \frac{1}{2}\sum_{\alpha,\beta,\delta,\sigma}\sum_{s_i}\bar{v}_\alpha(p_i, s_i)\gamma^0_{\alpha\beta}\sum_{s_f}v_\beta(p_f, s_f)\bar{v}_\delta(p_f, s_f)\gamma^0_{\delta\sigma}v_\sigma(p_i, s_i) \quad (5.52)$$

Using Eqn. 3.138 we can see that the sum over s_f will yield the $\beta\delta^{th}$ element of :

(LEFTOVER: Check to see if this double minus sign can be explained)

$$-\left(\frac{-\not{p}_f + m}{2m}\right)$$

using this we obtain:

$$-\frac{1}{2}\sum_{\alpha,\beta,\delta,\sigma}\sum_{s_i}\bar{v}_\alpha(p_i, s_i)\gamma^0_{\alpha\beta}\gamma^0_{\delta\sigma}\bar{v}_\sigma(p_i, s_i)(\left(\frac{-\not{p}_f + m}{2m}\right))_{\beta\delta}\gamma^0_{\delta\sigma}v_\sigma(p_i, s_i) \quad (5.53)$$

performing the now familiar operation and keeping track of the indices the reader can verify rather easily that we will obtain after summing over s_i :

$$\frac{1}{2}\sum_{\alpha\sigma}\left(\frac{-\not{p}_i+m}{2m}\right)_{\sigma\alpha}[\gamma^0(\left(\frac{-\not{p}_f+m}{2m}\right)\gamma^0]_{\alpha\sigma} \tag{5.54}$$

which is simply the $\frac{1}{2}$ the trace of the matrix (after summing over σ) or:

$$Tr[\frac{(-\not{p}_i+m)}{2m}\gamma^0\frac{(-\not{p}_f+m)}{2m}\gamma^0]=Tr[\frac{(\not{p}_i-m)}{2m}\gamma^0\frac{(\not{p}_f-m)}{2m}\gamma^0] \tag{5.55}$$

Using the cyclical property of the trace, recalling that products of odd numbers of γ matrices have zero trace we can easily recognize that the result of this operation will equal to Eqn. 5.48 since the coefficient of m^2 is greater than zero. We have deliberately written the Eqn. 5.55 in the two forms to be able to use a procedure in which we can reverse the momentum direction in other processes to obtain positron cross sections from electron cross sections by this process that we will call crossing symmetry. The reversal of momentum is clearly seen in the propagator and plane wave exponential terms, here it has been introduced so that modulus squared spin summed and averaged amplitudes can be modified by simply substituting an overall negative momentum to obtain the positron cross section using the electron cross section.

Note that after writing down Eqn. 5.55 we can easily see that to first order we will see exactly the same differential cross section that we saw for electron scattering-which is not surprising. However it should be easy to show (by keeping track of the signs in the series for S_{fi}) that at second and higher orders the positron and electron cross sections will differ.

(LEFTOVER: This whole bit about positrons needs a revisit).

Chapter 6

Moller ($e^-e^- \to e^-e^-$) and Bhabha scattering ($e^+e^- \to e^+e^-$)

In this chapter, we will calculate the differential cross section for the scattering of two electrons in the ultrarelativistic limit. The result is valid for any two identical fermions. We will then use our description of positrons being negative energy electrons moving backward in time derive the differential cross section for the scattering of electrons with positrons again generalizable to the scattering of any fermion with its anti-fermion partner.

The calculation will be done in the center of mass frame with particles approaching each other with equal and opposite momenta. In the last leg of the calculation we will take the ultra-relativistic limit or $E >> m$ for each electron. This problem will be treated as one where one particle sees the potential of of the other as derived using the relation

$$\Box A^\mu(x) = 4\pi J^\mu(x) \tag{6.1}$$

first encountered in chapter 1. Its important to note that since the electron creating the potential is moving all four components of its current density (due to its electric charge and motion) will be present-in contrast to Mott scattering in the previous chapter where only A^0 arose since the nucleus was stationary.

We will insert the initial and final wave functions in the series for S_{fi} along with the $A^\mu(x)$ generated by one of the electrons. It will then be obvious that in the scattering *either* electron can be seen as the source of the 4-vector potential and we will then correct for the fact that electrons are indistinguishable and that any amplitude must go to zero when both are in the same quantum state (fermions). This will give rise to two contributing terms in the scattering amplitude. We will follow a similar sequence of steps, ie modulus squaring the amplitude, averaging and summing over the initial and final spin states, counting the final number density of states and dividing by an incident flux and obtaining a differential cross section with respect to the exit variables of one of the electrons.

The two particle initial and final states will increase complexity of the calculation significantly however, the expressions for the flux and the spin averages and sums will be considerably more involved than in the previous chapter. We will continue to draw from results derived in the previous chapter.

Finally we will use the interpretation of negative energy electrons moving backward in time to represent positrons to obtain a result for $e^+e^- \rightarrow e^+e^-$ scattering as well and apply it to the electroproduction of muons via the process $e^+e^- \rightarrow \mu^+\mu^-$.

6.1 Moller Scattering: Setting up the scattering amplitude using the potential, initial and final wave functions

We will begin this section with the calculation of the 4-vector potential that we will use in our perturbation series. As mentioned we consider one electron as "seeing" the A^μ (potential) due to the J^μ of the other (current) before considering them as identical Fermions correcting for that. The potential due to a current is given by the co-variant equation (Chapter 1. Eqn. 1.76):

$$\Box A^\mu(x) = 4\pi J^\mu(x) \tag{6.2}$$

After solving for $A^\mu(x)$ we will simply use it in the perturbative series defined in equation 4.85 (4.86 for e^- backward in time) in a similar manner to what was done in Chapter 5 for Mott Scattering.

We begin by first making a choice for the current density. We know what the 4-component charge current density J^μ created by an electron propagating a series of space-time points y is: $e\bar{\psi}(y)\gamma^\mu\psi(y)$ (see the discussion preceding Eqn. 3.58) where the e is the electron's charge. However we want to derive the potential *during* the scattering process and so we choose based on intuition, the following choice which incorporates an initial and a final state :

$$J^\mu(y) = e\bar{\psi}_f(y)\gamma^\mu\psi_i(y) \tag{6.3}$$

We will see later that this choice is physically justified. To solve Eqn. 6.2, we claim that if there exists a function $D_F(x-y)$ such that $\Box D_F(x-y) =$

$4\pi\delta^4(x-y)$ then, $A^\mu(x) = \int d^4y D_F(x-y)J^\mu(y)$. This solution for $A^\mu(x)$ is telling us that the trajectory of a charged particle through one region of space and time is related to the potential in another such region. Note y is a dummy variable and is integrated over.

It is easy to verify the relation :

$$\Box A^\mu(x) = \Box \int d^4y D_F(x-y)J^\mu(y) \tag{6.4}$$

where the d'Alembertian $\Box = \Box_x$ is in terms of $x^\mu = (x_0, \vec{x})$. By definition of $D_F(x-y)$

$$\Box A^\mu(x) = \int d^4y 4\pi\delta^4(x-y)J^\mu(y) \tag{6.5}$$

which by definition of the Dirac δ function $f(x) = \int d^4y\delta^4(x-y)f(y)$ gives

$$\Box A^\mu(x) = 4\pi J^\mu(x) \tag{6.6}$$

Next we attempt to find a $D_F(x-y)$ using the Fourier transform technique.

Let $D_F(x-y)$ have the following representation:

$$D_F(x-y) = \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} D_F(q) \tag{6.7}$$

Applying the operator $\Box_x$ (D'Alembertian with respect to x) to both sides of the equation:

$$\Box_x D_F(x-y) = -q^2 \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} D_F(q) \tag{6.8}$$

By definition of $D_F(x-y)$, the left hand sign is just proportional the Dirac δ function:

$$4\pi\delta^4(x-y) = -q^2 \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} D_F(q) \tag{6.9}$$

By using the well known representation of the Dirac δ function then:

$$\delta^4(x-y) = \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} \tag{6.10}$$

we can rewrite Eqn. 6.9 to give

$$4\pi \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} = -q^2 \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} D_F(q) \tag{6.11}$$

Removing the $\int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)}$ from both sides we obtain:

$$D_F(q) = -\frac{4\pi}{q^2} \tag{6.12}$$

with a solution for $A^\mu(x)$:

$$A^\mu(x) = \int d^4y \int \frac{d^4q}{(2\pi)^4} \left(-\frac{4\pi}{q^2}\right) e^{-iq\cdot(x-y)} J^\mu(y) \tag{6.13}$$

We will now insert this potential into our series for the scattering amplitude. Note that to first order, the series for the scattering amplitude is:

$$\begin{aligned} S_{fi} = \delta_{fi} - ie \int d^4x_1 \bar{\psi}_f(x_1) \not{A}(x_1) \psi_i(x_1) \\ - ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2) \not{A}(x_2) S_F(x_2 - x_1) \not{A}(x_1) \psi_i(x_1) + \cdots \end{aligned} \tag{6.14}$$

We will expand the series for S_{fi} to order e only. Using x instead of x_1 as our dummy variable and inserting the solution for $A^\mu(x_1)$ into the first order term in Eqn. 6.14 we obtain:

$$S_{fi} = -ie \int d^4x \int d^4y \int \frac{d^4q}{(2\pi)^4} \bar{\psi}_f \gamma^\mu \psi_i(x) \left(-\frac{4\pi}{q^2}\right) e^{-iq\cdot(x-y)} J_\mu(y) \tag{6.15}$$

(we ignore the δ_{fi} in (6.14) because we're only interested in the amplitude for scattering to happen).

Note that expression (6.15) would just as well apply to either electron in the scattering process. This will become more apparent in the following lines. As in Chapter 5 we assume that the whole scattering process takes place in a large box of side L and volume $V = L^3$ over a time T, both the space and the time over which the scattering is observed is much greater than those for the interaction itself.

We label the momentum and spin variables of the incoming electrons by 1 and 2 and of the outgoing electrons by $1'$ and $2'$ and will normalize them appropriately. We will begin to treat the problem in the following way: electron 1 "sees" the potential created by the 4-current of electron 2. We recall that the probability current for a particle is $\bar{\psi}\gamma^\mu\psi$ as we calculated in Eqn. 3.55. Thus we assume that this times the charge $e \times \bar{\psi}\gamma^\mu\psi$ gives the charged current.

We begin by writing the following initial state and final state wave functions for electron 1 which we consider as scattering in the potential created by electron 2,

$$\begin{aligned}\psi_1(x) &= \sqrt{\frac{m}{E_1 V}} e^{-i\cdot p_1 \cdot x} u(p_1, s_1) \\ \bar{\psi}_{1'}(x) &= \sqrt{\frac{m}{E_{1'} V}} e^{+\cdot p_{1'} \cdot x} \bar{u}(p_{1'}, s_{1'})\end{aligned} \tag{6.16}$$

For electron 2 whose wave functions will be used to compute $J^\mu(y)$ (which gives us $A^\mu(x)$ via Eqn. 6.3 for use in Eqn. 6.13):

$$\begin{aligned}\psi_2(y) &= \sqrt{\frac{m}{E_2 V}} e^{-i\cdot p_2 \cdot y} u(p_2, s_2) \\ \bar{\psi}_{2'}(y) &= \sqrt{\frac{m}{E_{2'} V}} e^{+i\cdot p_{2'} \cdot y} \bar{u}(p_{2'}, s_{2'})\end{aligned} \tag{6.17}$$

We now adopt a short hand for the spinors in expressions 6.16 and 6.17, so $u(p_i, s_i) = u_i$ thus for example $u(p_2, s_2) = u_2$ and $\bar{u}(p_{2'}, s_{2'}) = \bar{u}_{2'}$ etc. Using $J^\mu(y) = e \times \bar{\psi}_{2'}(y)\gamma^\mu\psi_2(y)$ in Eqn. 6.13, we obtain for S_{fi} after collecting all the exponentials with some simplification:

$$S_{fi} = +ie^2 \int d^4x \int d^4y \int \frac{d^4q}{(2\pi)^4} \sqrt{\frac{m}{E_{1'}V}} \sqrt{\frac{m}{E_{2'}V}} \sqrt{\frac{m}{E_1 V}} \sqrt{\frac{m}{E_2 V}} e^{-i(p_2 y + p_{2'}\cdot y - q)\cdot y} \tag{6.18}$$

$$\times \left(\frac{4\pi}{q^2}\right) e^{-i(p_1 - p_{1'} + q)\cdot x} \times \bar{u}_{1'}\gamma_\mu u_1 \times \bar{u}_{2'}\gamma^\mu u_2$$

using the definitions for Dirac δ functions we have

$$\begin{aligned}\int d^4x e^{i(p_{1'} - p_1 - q)\cdot x} &= (2\pi)^4 \delta^4(p_{1'} - p_1 - q) \\ \int d^4y e^{i(p_{2'} - p_2 + q)\cdot y} &= (2\pi)^4 \delta^4(p_{2'} + q - p_2)\end{aligned} \tag{6.19}$$

and so our expression for S_{fi} becomes:

$$S_{fi} = ie^2(2\pi)^4(2\pi)^4\frac{m^2}{V^2\sqrt{E_{1'}E_{2'}E_1E_2}}\int\frac{d^4q}{(2\pi)^4}\delta^4(p_{1'}-p_1-q)\delta^4(p_{2'}-p_2+q) \tag{6.20}$$

$$\times\bar{u}_{1'}\gamma_\mu u_1\frac{4\pi}{q^2}\bar{u}_{2'}\gamma^\mu u_2$$

showing that the 4-momentum given up by one electron (q) is gained. Our next step is to simply integrate over $\int d^4q\delta^4(p_{2'}-p_2+q)$. This will set $q=(p_2-p_{2'})$ in the remainder of the expression including the remaining δ function, simplifying we get:

$$S_{fi} = \frac{ie^2m^2(2\pi)^4}{V^2\sqrt{E_{1'}E_1E_{2'}E_2}}\delta^4(p_{1'}+p_{2'}-p_1-p_2)\times\bar{u}_{2'}\gamma^\mu u_2\frac{4\pi}{(p_2-p_{2'})^2}\times\bar{u}_{1'}\gamma_\mu u_1 \tag{6.21}$$

We now have an expression with one "overall" Dirac delta function expressing the conservation of 4-momentum, in simplifying S_{fi} the 4-vector potential $A^\mu(x)$ has yielded $\frac{4\pi}{q^2}$ via its' Fourier transform, we see it play the role of mediating the transfer of momentum between an initial and a final particle state. The function $\frac{4\pi}{q^2}$ is the momentum space representation of A^μ and if written in terms of co-ordinates x, we would have called it a co-ordinate space representation, note that $q=p_2-p_{2'}$. (WHAT IS PHYSICALLY THE FOURIER TRANSFORM OF THE 4-POTENTIAL??)

Note that by looking at Eqn. 6.21 it is easy to see that one can consider this scattering as electron 1 interacting with the field created by electron 2 or vice-versa-one cannot say that one electron "feels" the potential of the other. The expression in equation 6.21 carries no information on which particle is doing the scattering and which is providing the field. This is the physical justification for our choice of transition current in Eqn. 6.3.

We now consider the fact that we are dealing with identical fermions. If we observe an electron in a final state we have no way of determining whether it is electron (1) or electron (2). One way to do this would be to replace 1 by 2 in the amplitude we have just calculated and add this amplitude to the previous however this would not take into account the Pauli exclusion principle. With a little bit of thought we see that the following amplitude satisfies the Pauli exclusion principle:

$$S_{fi} = \frac{ie^2m^2(2\pi)^4}{V^2\sqrt{E_{1'}E_1E_{2'}E_2}}\delta^4(p_{1'}+p_{2'}-p_1-p_2) \tag{6.22}$$

$$\times \left\{ \bar{u}_{2'}\gamma^\mu u_2 \frac{4\pi}{(p_2 - p_{2'})^2} \bar{u}_{1'}\gamma_\mu u_1 - \bar{u}_{2'}\gamma^\alpha u_1 \frac{4\pi}{(p_1 - p_{2'})^2} \bar{u}_{1'}\gamma_\alpha u_2 \right\}$$

If we set $1 = 2$, or $1' = 2'$ we will get zero for the scattering amplitude. Thus in the final or initial two-particle state the probability for them to be in the same quantum state, (specified by labels 1,2, for the initial and $1'$ and $2'$, for the final states) is zero. The next step is to modulus square the amplitude, insert counting factors and account for the fact that we dont observe the initial and final spins of the scattering particles. We refer to the momentum transfer $p_2 - p_{2'}$ as q_1 and to the momentum transfer $p_1 - p_{2'}$ as q_2.

6.2 Moller Scattering: Converting $| S_{fi} |^2$ into a probability inserting the counting factors

(LEFTOVER: above, is this really a probability ?)

Now we must multiply S_{fi} by its complex conjugate and insert the counting factors for *two* exiting particles and then finally insert the flux factor (next section).

We rewrite S_{fi} in the following way

$$S_{fi} = \frac{m^2(2\pi)^4}{V^2\sqrt{E_1 E_{1'} E_2 E_{2'}}}\delta^4(p_{1'} + p_{2'} - p_1 - p_2)M_{fi} \tag{6.23}$$

where of course M_{fi} is given by:

$$M_{fi} = ie^2 4\pi(\bar{u}_{2'}\gamma^\mu u_2 \frac{g_{\mu\nu}}{(p_2 - p_{2'})^2} \bar{u}_{1'}\gamma^\nu u_1 - \bar{u}_{2'}\gamma^\alpha u_1 \frac{g_{\alpha\beta}}{(p_1 - p_{2'})^2} \bar{u}_{1'}\gamma_\beta u_2) \tag{6.24}$$

where the reader is reminded that the factor of 4π is associated with each appearance of the metric tensor. We can write down the second expression in Eqn. 6.24 simply by looking at the first term, and in fact for any process involving a fermion (and so scattering forward in time) scattering off another we can generalize and state that the following factors written from left to right will represent the amplitude to <u>first order (!)</u>, (anti-fermions are treated in the section on Bhabha scattering.

1. Incoming electrons $\to u(p, s)$
2. Outgoing electrons $\to \bar{u}(p', s')$
3. A factor of $ie\gamma^\mu$ representing the coupling to the virtual photon and a factor of $\frac{4\pi g_{\mu\nu}}{(p-p')^2}$ representing the virtual photon.

Pictorially we can depict the two processes contributing to the amplitude in figures 6.1, all the factors above are written in and is easy to see that with the correct symmetry factor (simply considering the Pauli principle and attaching a factor -1 to the second process as discussed) applied one gets the amplitude in equation 6.24.

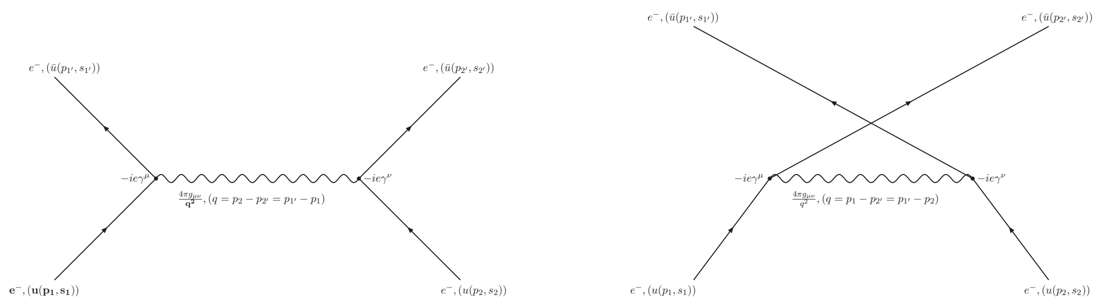

Fig. 6.1 Graphical depiction of $e^-e^- \to e^-e^-$. The process on the left has exchanged momentum $q_1 = p_{1'} - p_1 = p_2 - p_{2'}$, the exchange amplitude (on the right) has momentum transfer $q_2 = p_1 - p_{2'} = p_{2'} - p_1$ (one either exchanges the primed labels $1', 2'$ or the unprimed labels 1,2)

Returning to the S matrix element we have

$$\mid S_{fi} \mid^2 = \frac{m^4}{V^4 E_1 E_2 E_{1'} E_{2'}} ((2\pi)^4 \delta^4(p_{1'} + p_{2'} - p_1 - p_2))^2 \mid M_{fi} \mid^2 \quad (6.25)$$

In analogy with the discussion preceding Eqn. 5.16, we must multiply the expression for $\mid S_{fi} \mid^2$ by a counting factor for *each* exiting electron. For the two electrons, these are: $\frac{Vd^3p_{1'}}{(2\pi)^3}$ and $\frac{Vd^3p_{2'}}{(2\pi)^3}$. This will cancel two powers of the volume (V^2) in the denominator using $dP = |S_{fi}|^2 \frac{Vd^3p_{1'}}{(2\pi)^2} \frac{Vd^3p_{2'}}{(2\pi)^2}$ (see development before and including Eqn. 5.16)

$$dP = \frac{m^4}{V^2 E_1 E_2}((2\pi)^4 \delta^4(p_{1'} + p_{2'} - p_1 - p_2))^2 \mid M_{fi} \mid^2 \frac{d^3p_{1'}}{(2\pi)^3 E_{1'}} \frac{d^3p_{2'}}{(2\pi)^3 E_{2'}} \tag{6.26}$$

where $E_{1'}$ and $E_{2'}$ have been moved to the right under $d^3p_{1'}$ and $d^3p_{2'}$.

Now we deal with the square of the Dirac delta function in Eqn. 6.26. In Chapter 5 we had written down an expression for the result of squaring a Dirac delta function and so using the result in Eqn. 5.22

$$((2\pi)^4 \delta^4(p_{1'} + p_{2'} - p_1 - p_2))^2 = (2\pi)^4 VT \delta^4(p_{1'} + p_{2'} - p_1 - p_2)$$

where V is the volume of the box and T the (large!) time in which the scattering takes place. Inserting this, another power of volume cancels and dividing by T, we are left with $\frac{dP}{T}$ on the left which we had interpreted as a rate in Chapter 5, Eqn. 5.25

$$dR = \frac{m^4}{VE_1E_2}(2\pi)^4 \delta^4(p_{1'} + p_{2'} - p_1 - p_2)^2 \mid M_{fi} \mid^2 \frac{d^3p_{1'}}{(2\pi)^3 E_{1'}} \frac{d^3p_{2'}}{(2\pi)^3 E_{2'}} \tag{6.27}$$

6.3 Moller scattering: The two-body incident flux and getting a cross section

We know that the rate of interactions is $d\sigma$ multiplied by the flux of incoming particles; $dR = \mid \vec{j}_i \mid d\sigma$. The flux in this case is of two incoming particles. Recall we used the flux of a single incoming particle in Chapter 5 (see discussion preceding Eqn. 5.34). We will therefore use

$$\mid \vec{j}_i \mid = \frac{|\vec{v}_1 - \vec{v}_2|}{V}$$

for the case of two beams of electrons approaching head-on. LEFTOVER: GENERALIZE IN EVERY PL POSSIBLE TO JUSTIFY FEYNMANN RULES.

At this point, we make a slight diversion to calculate $\mid \vec{j}_i \mid$ explicitly. We note that $\frac{|\vec{v}_1 - \vec{v}_2|}{V} = \left| \frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2} \right| \times \frac{1}{V}$ and ask the reader to consider

$E_1E_2 \times \left|\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2}\right| \times \frac{1}{V}$. We do this because Eqn. 6.23 contains a factor of $\frac{1}{\sqrt{E_1E_2}}$ which when squared along with everything else will result in a factor of

$$E_1E_2\left|\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2}\right| \times \frac{1}{V} \tag{6.28}$$

in the denominator of the expression after division by $\mid \vec{j}_i \mid$. Back to (6.28):

$$E_1E_2\left|\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2}\right| = E_1E_2\left(\frac{|\vec{p}_1|^2}{E_1^2} - \frac{2\vec{p}_1\cdot\vec{p}_2}{(E_1)(E_2)} + \frac{|\vec{p}_2|^2}{E_2^2}\right)^{\frac{1}{2}} \tag{6.29}$$

$$= \left(|\vec{p}_1|^2(E_2)^2 - 2\vec{p}_1\cdot\vec{p}_2(E_1E_2) + |\vec{p}_2|^2(E_1)^2\right)^{\frac{1}{2}}$$

using different masses for a general result one may need in the future we use $\mid \vec{p}_{1,2} \mid^2 = E_{1,2}^2 - m_{1,2}^2$ to obtain:

$$= \left((E_1^2 - m_1^2)(E_2)^2 - 2\vec{p}_1\cdot\vec{p}_2(E_1E_2) + (E_2^2 - m_2^2)(E_1)^2\right)^{\frac{1}{2}}$$

$$= \left(2E_1^2E_2^2 - 2\vec{p}_1\cdot\vec{p}_2(E_1E_2) - m_1^2E_2^2 - m_2^2E_1^2\right)^{\frac{1}{2}} \tag{6.30}$$

We now consider the Lorentz scalar:

$$(p_1\cdot p_2)^2 = (E_1E_2 - \vec{p}_1\cdot\vec{p}_2)^2 \tag{6.31}$$

Expanding out Eqn 6.31 by using $|\vec{p}_{1,2}|^2 = E_{1,2}^2 - m_{1,2}^2$

$$\begin{aligned}(E_1E_2 - \vec{p}_1\cdot\vec{p}_2)^2 &= E_1^2E_2^2 - 2E_1E_2\vec{p}_1\cdot\vec{p}_2 + E_1^2E_2^2 - m_2^2E_1^2 - m_1^2E_2^2 + m_1^2m_2^2 \\ &= 2E_1^2E_2^2 - 2\vec{p}_1\cdot\vec{p}_2(E_1E_2) - m_1^2E_2^2 - m_2^2E_1^2 + m_1^2m_2^2 \\ &= (p_1\cdot p_2)^2\end{aligned} \tag{6.32}$$

By comparing Eqns. 6.30 and 6.32 we obtain the result:

$$E_1E_2\left|\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2}\right| = \sqrt{(p_1\cdot p_2)^2 - m_1^2m_2^2} \tag{6.33}$$

Of course for our example we have two electrons and $m = m_1 = m_2$, doing this we get:

$$E_1 E_2 \left| \frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2} \right| = \sqrt{(p_1 \cdot p_2)^2 - m^4} \tag{6.34}$$

The reader is reminded that in Eqn. 6.27 the factor of $\frac{1}{\sqrt{E_1 E_2}}$ can be absorbed into writing the flux factor. Note also that $\sqrt{(p_1 \cdot p_2)^2 - m_1 m_2}$ is Lorentz invariant. Putting in all the counting and flux factors, we obtain the following expression for the differential cross section after all factors of V cancel:

$$d\sigma = \frac{(2\pi)^4 \delta^4 (p_1 + p_2 - p_{1'} - p_{2'})}{\sqrt{(p_1 \cdot p_2)^2 - m^4}} |M_{fi}|^2 \times m^2 \times \frac{m}{E_{1'}} \frac{d^3 p_{1'}}{(2\pi)^3} \frac{m}{E_{2'}} \frac{d^3 p_{2'}}{(2\pi)^3} \tag{6.35}$$

where M_{fi} is defined in Eqn. (6.24).

6.4 Moller Scattering: Averaging and summing over the initial and final spins, Casimir's trick again

We will now average $|M_{fi}|$ over the initial and sum over the final spins of the two electrons.

$$\begin{aligned} |M_{fi}|^2 = e^4 (4\pi)^2 \Big(& \frac{\bar{u}_{2'} \gamma^\mu u_2 \bar{u}_{1'} \gamma_\mu u_1}{q_1^2} - \frac{\bar{u}_{2'} \gamma^\alpha u_1 \bar{u}_{1'} \gamma_\alpha u_2}{q_2^2} \Big) \\ \times \Big(& \frac{\bar{u}_{2'} \gamma^\nu u_2 \bar{u}_{1'} \gamma_\nu u_1}{q_1^2} - \frac{\bar{u}_{2'} \gamma^\beta u_1 \bar{u}_{1'} \gamma_\beta u_2}{q_2^2} \Big)^* \end{aligned} \tag{6.36}$$

Where $q_1 = p_2 - p_{p'}$ and $q_2 = p_1 - p_{2'}$. Recall that in Chapter 5 we had derived the result $(\bar{u}_1 \gamma^\alpha u_2)^* = (\bar{u}_2 \gamma^\alpha u_1)$ (equation 5.46). We use this in Eqn 6.36 and introduce explicit spinor indices, i.e $\bar{u}_1 \gamma^\alpha u_2 = \bar{u}_{1,\sigma} \gamma^\alpha_{\sigma\rho} u_{2,\rho}$, etc in the next few steps.

We first expand Eqn. 6.36 and use different indices for the γ matrices in the complex conjugate term since the term and its complex conjugate are two different sums.

$$
\begin{aligned}
|M_{fi}|^2 = e^4(4\pi)^2(&\frac{\bar{u}_{2'}\gamma^\mu u_2 \bar{u}_{1'}\gamma_\mu u_1 \bar{u}_1 \gamma_\nu u_{1'} \bar{u}_2 \gamma^\nu u_{2'}}{q_1^4} \\
&+ \frac{\bar{u}_{2'}\gamma^\alpha u_1 \bar{u}_{1'}\gamma_\alpha u_2 \bar{u}_2 \gamma_\beta u_{1'} \bar{u}_1 \gamma^\beta u_{2'}}{q_2^4} \\
&- \frac{\bar{u}_{2'}\gamma^\mu u_2 \bar{u}_{1'}\gamma_\mu u_1 \bar{u}_2 \gamma_\alpha u_{1'} \bar{u}_1 \gamma^\alpha u_{2'}}{q_1^2 q_2^2} \\
&- \frac{\bar{u}_{2'}\gamma^\alpha u_1 \bar{u}_{1'}\gamma_\alpha u_2 \bar{u}_1 \gamma_\mu u_{1'} \bar{u}_2 \gamma^\mu u_{2'}}{q_1^2 q_2^2})
\end{aligned}
\tag{6.37}
$$

Although there are 4 terms in Eqn. 6.37, we need only consider two of these. Taking only the spinors in the first term :

$$\frac{\bar{u}_{2'}\gamma^\mu u_2 \bar{u}_{1'}\gamma_\mu u_1 \bar{u}_1 \gamma_\nu u_{1'} u_2 \gamma^\nu u_{2'}}{q_1^4}$$

and changing $2 \leftrightarrow 1$ we obtain

$$\frac{\bar{u}_{2'}\gamma^\mu u_1 \bar{u}_{1'}\gamma_\mu u_2 \bar{u}_2 \gamma_\nu u_{1'} \bar{u}_1 \gamma^\nu u_{2'}}{q_2^4}$$

This is simply the second term of Eqn. 6.37. Note the indices of the gamma matrices run from 0-3 if they are free (thus the μ and ν can be replaced with α and β), so the second term can be obtained from the first by $2 \leftrightarrow 1$. The same holds true for the third and fourth terms as well–the fourth can be obtained from the third by $2 \leftrightarrow 1$–the reader is urged to check this.

Lets begin with the first term. We introduce indices explicitly, and recognize that terms like $\bar{u}_2\gamma^\nu u_{2'}$ are scalars and can be moved around and it is most convinient to put a u_i just before a $\bar{u}_i$ (same index) in preparation for utilizing the relation $\sum_s u_i\bar{u}_i = \frac{\not{p}+m}{2m}$. We begin by re-writing the first term in curly brackets and rearranging a bit:

$$\frac{1}{q_1^4}\bar{u}'_{2\rho}\gamma^\mu_{\rho\tau}u_{2\tau}\bar{u}_{2\epsilon}\gamma^\nu_{\epsilon\eta}u'_{2\eta}\bar{u}'_{1\sigma}\gamma_{\mu,\sigma\nu}u_{1\nu}\bar{u}_{1\delta}\gamma_{\nu,\delta\chi}u'_{1\chi} \tag{6.38}$$

To sum over the spins, we pick out the specific element of a spinor in Eqn (6.38) for one particular spin. As an example we can see clearly that this means $\sum_{s_1 s_2} u_{2\tau}\bar{u}_{2\epsilon} = (\frac{\not{p}_2+m}{2m})_{\tau\epsilon}$.

Using this, we simplify Eqn. 6.38 to the following expressions after summing over the two possible spins of all spinors that appear as $u_i\bar{u}_i$ (thus electron 1, and 2):

$$\frac{1}{q_1^4}\bar{u}'_{2\rho}\gamma^{\mu}_{\rho\tau}\left(\frac{\not{p}_2+m}{2m}\right)_{\tau\epsilon}\gamma^{\nu}_{\epsilon\eta}u'_{2\eta}\bar{u}'_{1\sigma}\gamma_{\mu,\sigma\nu}\left(\frac{\not{p}_1+m}{2m}\right)_{\nu\delta}\gamma_{\nu,\delta\chi}u'_{1\chi} \tag{6.39}$$

Note that by keeping track of the indices of the γ matrices and of the spinors we can write:

$$\frac{1}{q_1^4}\bar{u}'_{2\rho}\left(\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma^{\nu}\right)_{\rho\eta}u'_{2\eta}\bar{u}_{1\sigma}\left(\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma_{\nu}\right)_{\sigma\chi}u'_{1\chi} \tag{6.40}$$

Since all of the indices are written explicitly in the above Eqn. 6.40, we can move them around and sum over spins thus we obtain:

$$\frac{1}{q_1^4}u'_{2\eta}\bar{u}'_{2\rho}\left(\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma^{\nu}\right)_{\rho\eta}u'_{1\chi}\bar{u}'_{1\sigma}\left(\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma_{\nu}\right)_{\sigma\chi}$$

Inserting the spin sums over the scattered electrons $1'$ and $2'$, we obtain:

$$\frac{1}{q_1^4}\left(\frac{\not{p}_{2'}+m}{2m}\right)_{\eta\rho}\left(\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma^{\nu}\right)_{\rho\eta}\left(\frac{\not{p}_{1'}+m}{2m}\right)_{\chi\sigma}\left(\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma_{\nu}\right)_{\sigma\chi} \tag{6.41}$$

Summing over ρ and σ in Eqn 6.41, we obtain

$$\frac{1}{q_1^4}\left(\frac{\not{p}_{2'}+m}{2m}\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma^{\nu}\right)_{\eta\eta}\left(\frac{\not{p}_{1'}+m}{2m}\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma_{\nu}\right)_{\chi\chi} \tag{6.42}$$

We note that the repeated indices $\eta\eta$ and $\chi\chi$ signify a trace of the matrices that appear within the brackets; η and χ are distinct indices meaning we have product of two traces. We multiply by $\frac{1}{4}$ for to average over the initial spins and write:

$$\frac{1}{64m^4q_1^4}\times Tr\left[(\not{p}_{2'}+m)\gamma^{\mu}(\not{p}_2+m)\gamma^{\nu}\right]Tr\left[(\not{p}_{1'}+m)\gamma_{\mu}(\not{p}_1+m)\gamma_{\nu}\right] \tag{6.43}$$

We now turn to the third (mixed) term

$$-\frac{\bar{u}_{2'}\gamma^{\mu}u_2\bar{u}_{1'}\gamma_{\mu}u_1\bar{u}_2\gamma_{\alpha}u_{1'}\bar{u}_1\gamma^{\alpha}u_{2'}}{q_1^2q_2^2}$$

Using the fact that terms like $\bar{u}'_j\gamma^{\alpha}u_i$, $\bar{u}_j\gamma^{\alpha}u'_i$, etc are scalars, we move $\bar{u}_2\gamma_{\alpha}u_{1'}$ to a location in the product where we will be able to recognize three spin sums that we can carry out:

$$-\frac{\bar{u}_{2'}\gamma^{\mu}u_2\bar{u}_2\gamma_{\alpha}u_{1'}\bar{u}_{1'}\gamma_{\mu}u_1\bar{u}_1\gamma^{\alpha}u_{2'}}{q_1^2q_2^2} \tag{6.44}$$

Introducing explicit indices and summing over spins 2, $1'$ and 1:

$$-\frac{1}{q_1^2q_2^2}\bar{u}'_{2\rho}\left(\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma_{\alpha}\frac{\not{p}_{1'}+m}{2m}\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma^{\alpha}\right)_{\rho\delta}u'_{2\delta} \tag{6.45}$$

rearranging:

$$-\frac{1}{q_1^2q_2^2}u'_{2\delta}\bar{u}'_{2\rho}\left(\gamma^{\mu}\frac{\not{p}_2+m}{2m}\gamma_{\alpha}\frac{\not{p}_{1'}+m}{2m}\gamma_{\mu}\frac{\not{p}_1+m}{2m}\gamma^{\alpha}\right)_{\rho\delta} \tag{6.46}$$

After summing over spin $2'$ and multiplying by $\frac{1}{4}$ for the average of the spins:

$$\frac{-1}{64m^4q_1^2q_2^2}\times Tr\left[(\not{p}_{2'}+m)\gamma^{\mu}(\not{p}_2+m)\gamma_{\alpha}(\not{p}_1'+m)\gamma_{\mu}(\not{p}_1+m)\gamma^{\alpha}\right] \tag{6.47}$$

We now have the traces in Eqns. 6.47 and 6.43. The reader is reminded that we intend to take the ultra-relativistic limit, ie all energies and momenta are taken to be much greater than the electron mass m, so this is set to zero in the numerator of both Eqn. 6.47 and Eqn. 6.43. In this limit the expressions we will manipulate are

$$\frac{1}{64m^4q_1^4}\times Tr\left[\not{p}_{2'}\gamma^{\mu}\,\not{p}_2\gamma^{\nu}\right]Tr\left[\not{p}_{1'}\gamma_{\mu}\,\not{p}_1\gamma_{\nu}\right] \tag{6.48}$$

and

$$\frac{-1}{64m^4q_1^2q_2^2}\times Tr\left[\not{p}_{2'}\gamma^{\mu}\,\not{p}_2\gamma_{\alpha}\,\not{p}_1'\gamma_{\mu}\,\not{p}_1\gamma^{\alpha}\right] \tag{6.49}$$

with $q_1 = p_{1'} - p_1$ and $q_2 = p_{2'} - p_1$.

In order to evaluate the traces in Eqns. 6.48-6.49 above, we will need to prove some more trace theorems. Recall that the last set of Theorems we proved in Chapter 5 were numbered until III. We now proove Theorems IV (a)-(d) and Theorem V (a) & (b) below:

Theorem IV (a) $\gamma_{\mu}\gamma^{\mu} = 4\mathbb{I}$

Proof:

$$\gamma_{\mu}\gamma^{\mu} = (\gamma_0, -\vec{\gamma})\cdot(\gamma_0, \vec{\gamma}) = \gamma_0\gamma^0 - (\gamma^1)^2 - (\gamma^2)^2 - (\gamma^3)^2$$

We know that $(\gamma^0)^2 = \mathbb{I}$ *and that the square of the rest of the* γs *is* $-\mathbb{I}$. *So trivially* $\gamma_\mu\gamma^\mu = 4\mathbb{I}$ *QED.*

Theorem IV (b) $\gamma_\mu \not{a}\gamma^\mu = -2\not{a}$.

Proof:

$$\gamma_\mu \not{a}\gamma^\mu = \gamma_\mu\gamma^\nu a_\nu\gamma^\mu = \gamma_\mu a_\nu\gamma^\nu\gamma^\mu$$

$$= \gamma_\mu a_\nu(2g^{\mu\nu} - \gamma^\mu\gamma^\nu) = \gamma_\mu a_\nu 2g^{\mu\nu} - a_\nu\gamma_\mu\gamma^\mu\gamma^\nu$$

$$= 2\not{a} - a_\nu 4\gamma^\nu = 2\not{a} - 4\not{a} = -2\not{a}$$

QED.

Theorem IV (c) $\gamma_\mu \not{a}\not{b}\gamma^\mu = 4a \cdot b\mathbb{I}$

Proof:

$$\gamma_\mu \not{a}\not{b}\gamma^\mu = a_\nu b_\alpha(\gamma_\mu\gamma^\nu\gamma^\alpha\gamma^\mu)$$

$$= a_\nu b_\alpha\gamma_\mu\gamma^\nu(2g^{\alpha\mu} - \gamma^\mu\gamma^\alpha) = a_\nu b_\alpha \cdot 2\gamma^\alpha\gamma^\nu - a_\nu b_\alpha\gamma_\mu\gamma^\nu\gamma^\mu\gamma^\alpha =$$

$$2\not{b}\not{a} + \gamma_\mu \not{a}\gamma^\mu \not{b}$$

using theorem IVb:

$$\gamma_\mu \not{a}\gamma^\mu = -2\not{a}$$

and we obtain

$$2\not{b}\not{a} - 2\not{a}\not{b} = 2a_\mu b_\nu\{\gamma^\mu, \gamma^\nu\}$$

which is $2a_\mu b_\nu \cdot 2g^{\mu\nu} = 4a \cdot b\mathbb{I}$ *QED.*

Theorem IV (d) $\gamma_\mu \not{a}\not{b}\not{c}\gamma^\mu = -2\not{c}\not{b}\not{a}$

Proof:

$$\gamma_\mu \not{a}\not{b}\not{c}\gamma^\mu = \gamma_\mu \not{a}\not{b}c_\nu\gamma^\nu\gamma^\mu$$

$$= \gamma_\mu \not{a}\not{b}c_\nu(2g^{\mu\nu} - \gamma^\mu\gamma^\nu) = 2\gamma_\mu \not{a}\not{b}c^\mu - \gamma_\mu \not{a}\not{b}\gamma^\mu \not{c}$$

the first term is $2\not{c}\not{a}\not{b}$ *the second by IV (c) is* $4a \cdot b\not{c}$: *we have* $2\not{c}\not{a}\not{b} - 4a \cdot b\not{c}$

Now $2\,\not{c}\not{a}\not{b} - 4a\cdot b\,\not{c} = 2\,\not{c}a_\mu b_\nu(\gamma^\mu\gamma^\nu) - 4a\cdot b\,\not{c}$ *using the anti-commutation relations we obtain:*

$$2\,\not{c}a_\mu b_\nu(2g^{\mu\nu} - \gamma^\nu\gamma^\mu) - 4a\cdot b\,\not{c}$$

which is $4a\cdot b\,\not{c} - 2\,\not{c}\not{b}\not{a} - 4a\cdot b\,\not{c} = -2\,\not{c}\not{b}\not{a}$ *QED.*

Theorem V (a) (See also Theorem II in Chapter 5) $Tr(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}$*. Proof:* $Tr(\gamma^\mu\gamma^\nu) = Tr(2g^{\mu\nu} - \gamma^\nu\gamma^\mu) = 2g^{\mu\nu}Tr(\mathbb{I}) - Tr(\gamma^\nu\gamma^\mu)$ *from the* γ *matrix anti-commutation relations and then using the cyclical property of the trace* $Tr(\gamma^\mu\gamma^\nu) = Tr(\gamma^\nu\gamma^\mu)$ *and re-arranging we find*

$$Tr(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}\,QED.$$

Theorem V (b) (See also Theorems III, Chapter 5) $Tr(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu) = 4g^{\beta\nu}g^{\alpha\mu} - 4g^{\mu\nu}g^{\alpha\beta} + 4g^{\mu\beta}g^{\alpha\nu}$*. Prooving this is rather straightforward, the plan is simpy to move* γ^ν *to the left of* γ^α *by using the anticommutation relations of the* γ *matrices followed by the cyclical property of the trace. Using* $\gamma^\beta\gamma^\nu = 2g^{\beta\nu} - \gamma^\nu\gamma^\beta$ *to move* γ^ν *to the left of* γ^β *and then again (with appropriate indices) to move past* $\gamma^\mu 4$ *and* γ^α *we obtain :*

$$Tr(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu) = 2g^{\beta\nu}Tr(\gamma^\alpha\gamma^\mu) - 2g^{\mu\nu}Tr(\gamma^\alpha\gamma^\beta) + 2g^{\mu\beta}Tr(\gamma^\alpha\gamma^\nu) - Tr(\gamma^\nu\gamma^\alpha\gamma^\mu\gamma^\beta)$$

By the cyclical property of the trace $Tr(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu) = Tr(\gamma^n\gamma^\alpha\gamma^\mu\gamma^\beta)$ *and using Theorem V (a) we obtain quite easily:*

$$Tr(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu) = 4g^{\beta\nu}g^{\alpha\mu} - 4g^{\mu\nu}g^{\alpha\beta} + 4g^{\mu\beta}g^{\alpha\nu}\,QED.$$

These are all the theorems we need. We now attack the middle "mixed" trace term in Eqn. 6.49. By the cyclicity of trace, we re-write this as (using only the trace term):

$$Tr(\gamma_\alpha\,\not{p}_{1'}\gamma_\mu\,\not{p}_1\gamma^\alpha\,\not{p}_{2'}\gamma^\mu\,\not{p}_2) \tag{6.50}$$

In Eqn 6.50 we use theorem IV (d) to replace $\gamma_\alpha\,\not{p}_{1'}\gamma_\mu\,\not{p}_1\gamma^\alpha$ with

$$-2\,\not{p}_1\gamma_\mu\,\not{p}_{1'}$$

to obtain:

$$-2Tr(\not{p}_1\gamma_\mu\,\not{p}_{1'}\,\not{p}_{2'}\gamma^\mu\,\not{p}_2) \tag{6.51}$$

In Eqn 6.51 we replace $\gamma_\mu\,\not{p}_{1'}\,\not{p}_{2'}\gamma^\mu$ with $4p_{1'}\cdot p_{2'}$ using Theorem IV (c) and Theorem II (Chapter 5) to obtain:

$$-8p_{1'} \cdot p_{2'} Tr(\not{p}_1 \not{p}_2) = -32(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) \tag{6.52}$$

Note, as discussed previously, the interchange $2 \leftrightarrow 1$ will give us an additional identical term, and so we write the total contribution to $\mid M_{fi} \mid^2_{SPIN}$ from just the trace as:

$$-64(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) \tag{6.53}$$

multiplying by the factors of $-\frac{e^4(4\pi)^2}{64q_1^2q_2^2}$

$$\frac{e^4(4\pi^2)}{q_1^2q_2^2}(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) \tag{6.54}$$

Next, we move on to evaluate the trace (dropping the $\frac{1}{m^4 64q_1^4}$ term for convenience for now) in Eqn 6.48 :

$$Tr(\not{p}_{2'}\gamma^\mu \not{p}_2\gamma^\nu)Tr(\not{p}_{1'}\gamma_\mu \not{p}_1\gamma_\nu) \tag{6.55}$$

$Tr(\not{p}_{2'}\gamma^\mu \not{p}_2\gamma^\nu)$. We note that this can be simply rewritten with only gamma matrices in the argument of the trace as:

$$Tr(\not{p}_{2'}\gamma^\mu \not{p}_2\gamma^\nu) = p'_{2,\alpha}p_{2,\beta}Tr(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu) \tag{6.56}$$

By Theorem V (b) we know this is simply:

$$\begin{aligned} &p'_{2,\alpha}p_{2,\beta}(4g^{\beta\nu}g^{\alpha\mu} - 4g^{\mu\nu}g^{\alpha\beta} + 4g^{\mu\beta}g^{\alpha\nu}) \\ &= 4p^\mu_{2'}p^\nu_2 - 4g^{\mu\nu}p_{2'} \cdot p_2 + 4p^\nu_{2'}p^\mu_2 \end{aligned} \tag{6.57}$$

The $Tr(\not{p}_{1'}\gamma_\mu \not{p}_1\gamma_\nu)$ term is similarly calculated to be :

$$4p_{1',\mu}p_{1,\nu} - 4p_{1'} \cdot p_1 g_{\mu\nu} + 4p_{1',\nu}p_{1,\mu} \tag{6.58}$$

And so Eqn 6.55 can be rewritten as:

$$16(p'_{1\mu}p_{1\nu} - (p_{1'} \cdot p_1)g_{\mu\nu} + p'_{1\nu}p_{1\mu})(p_2^{\mu\prime}p_2^\nu - p_{2'} \cdot p_2 g^{\mu\nu} + p_2^{\nu\prime}p_2^\mu) \tag{6.59}$$

Contracting all repeated indices ($g_{\mu\nu}g^{\mu\nu} = 4$) we expand the product in Eqn 6.59

$$p'_{1\mu}p_{1\nu}(-p_{2'} \cdot p_2 g^{\mu\nu}) = -(p_{1'} \cdot p_1)(p_{2'} \cdot p_2)$$

and it can be easily verified that this yields for Eqn. 6.59 after re-inserting $\frac{1}{m^4 64 q_1^4}$):

$$\frac{(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) + (p_1 \cdot p_{2'})(p_{1'} \cdot p_2)}{2(p_2 - p_{2'})^4} \tag{6.60}$$

At this point the reader is reminded that Eqn. 6.60 is the contribution from Eqn. 6.48 which in turn is the first term inside the brackets in Eqn. 6.37. By swapping $2 \leftrightarrow 1$ in Eqn. 6.60 we obtain the contribution from the second term inside the brackets in Eqn. 6.37

$$\frac{(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) + (p_2 \cdot p_{2'})(p_{1'} \cdot p_1)}{2(p_1 - p_{2'})^4} \tag{6.61}$$

Finally we gather the *total* contribution from the mixed term in Eqn. 6.54 and insert each contribution from each term within the brackets in Eqn. 6.37 we obtain for $|M_{fi}|^2$:

$$|M_{fi}|^2 = \frac{e^4(4\pi)^2}{2m^4} \times \left\{ \frac{(p_{1'} \cdot p_{2'})(p_1 \cdot p_2) + (p_1 \cdot p_{2'})(p_{1'} \cdot p_2)}{(p_1 - p_{1'})^4} + \frac{2(p_{1'} \cdot p_{2'})(p_1 \cdot p_2)}{(p_2 - p_{2'})^2 (p_1 - p_{2'})^2} \right. \tag{6.62}$$
$$\left. + \frac{(p_{2'} \cdot p_{1'})(p_1 \cdot p_2) + (p_{2'} \cdot p_2)(p_1 \cdot p_{1'})}{(p_{2'} - p_1)^4} \right\}$$

where $(p_{2'} - p_1)^4 = (p_2 - p_{1'})^4$ has been used.

In order to simplify things further and progress to the next step, we rewrite from Eqn 6.35:

$$d\sigma = m^4 (2\pi)^4 \frac{\delta^4(p_{1'} - p_1 + p_{2'} - p_2)}{\sqrt{(p_1 \cdot p_2)^2 - m^4}} \mid M_{fi} \mid^2_{SPIN} \frac{d^3 p_{1'}}{E_{1'}(2\pi)^3} \times \frac{d^3 p_{2'}}{E_{2'}(2\pi)^3} \tag{6.63}$$

At the beginning of this chapter, we decided to pick a simple set of kinematics with the incident electrons head on and in the center of mass. This means that $p_1 = (E, \vec{p})$ and $p_2 = (E, -\vec{p})$.

Thus the flux factor is simply:

$$\sqrt{(E^2 + |\vec{p}|^2)^2 - m^4} = \sqrt{(E^2 + E^2 - m^2)^2 - m^4} = \sqrt{4E^4 - 4E^2 m^2 + m^4 - m^4} \tag{6.64}$$

$$= 2\sqrt{E^2(E^2 - m^2)} = 2\sqrt{E^2|\vec{p}|^2} = 2E|\vec{p}|$$

We now need to think about integrating over the 4 dimensional Dirac δ function, in Eqn 6.63. Before proceeding we make a few comments about how we have chosen our kinematics. The incoming momenta are $p_1 = (E, \vec{p})$, $p_2 = (E, -\vec{p})$. The conservation of 4-momentum relates the final state momenta: $p_{1'} + p_{2'} = p_1 + p_2 \therefore \vec{p}_{2'} + \vec{p}_{1'} = 0$ and so we pick $\vec{p}_{1'} = \vec{p}'$ and $\vec{p}_{2'} = -\vec{p}'$. Note that this means that $E_{1'} = \sqrt{|\vec{p}| + m^2}$ is equal to $E_{2'}$. Thus $p_{1'} = (E', \vec{p}')$ and $p_{2'} = (E', -\vec{p}')$ finally since $p_{1'} + p_{2'} = p_1 + p_2$, so $E' = E$. Thus we conclude that $p_{1'} = (E, \vec{p}')$ and $p_{2'} = (E, -\vec{p}')$ with $|\vec{p}'| = |\vec{p}|$. We denote the angle between $\vec{p}$ and $\vec{p}'$ to be θ and will use the magnitudes of $\vec{p}$, $\vec{p}'$ at a convenient time during the calculation.

What remains now is to deal with the four dimensional δ function. In this problem we will assume one more thing: that we will observe one electron and integrate over the variables of the other. Let's integrate over electron (2) and promise to observe electron (1).

Now, in Eqn. 6.63, we have a 4-dimensional δ function but the differential over $d^3\vec{p}_{2'}$ is 3-dimensional. To tackle this dilemma we will show that

$$\frac{d^3\vec{p}_{2'}}{2E_{2'}} = \theta(p'_{20}) \cdot \delta(p_2'^2 - m^2) d^4 p_{2'} \tag{6.65}$$

After showing this, it becomes trivial to integrate $\delta^4(p_{1'} - p_1 + p_{2'} - p_2)$ since $d^4 p_{2'}$ is a 4-dimensional differential.

So lets check Eqn 6.65. $\theta(p'_{20})$ simply keeps $E_{2'} > 0$. Let's rewrite the right hand side of Eqn 6.65. The right hand side of Eqn 6.65 is:

$$\theta(p'_{20})\delta\left(\left[p'_{20} + \sqrt{|\vec{p}_{2'}|^2 + m^2}\right]\left[p'_{20} - \sqrt{|\vec{p}_{2'}|^2 + m^2}\right]\right) dp'_{20} d^3\vec{p}_{2'} \tag{6.66}$$

Note that $\int \delta[f(x)]dx = \sum_{i=1}^{N} \frac{1}{|f'(x_i)|}$ where $f'(x_i)$ is the derivative of $f(x)$ evaluated at the zeroes of $f(x)$. The zeroes of the argument of the δ function in Eqn. 6.66 are simply $\pm E_{2'}$, but $\theta(p'_{20})$ means we only consider $E_{2'}$. So proceeding:

$$\sum_{i=1}^{N} \frac{1}{f'(p'_{20})} = \frac{1}{2p'_{20}}|_{p'_{20}=E_{2'}} = \frac{1}{E_{2'}} \tag{6.67}$$

This gives us :

$$\theta(p'_{20}) \times \delta(p_2'^2 - m^2) d^4 p_{2'} = \frac{d^3 \vec{p'}_2}{2E_{2'}} \tag{6.68}$$

What we have shown in the previous analysis is that integrating over $\frac{d^3\vec{p}_{2'}}{2E_{2'}}$ is the same as integrating over $\theta(p'_{20})\delta(p_2'^2 - m^2)d^4p_{2'}$ and so we replace $\frac{d^3\vec{p}_{2'}}{E_{2'}(2\pi)^3}$ in Eqn. (6.63) by the left hand side of Eqn 6.68, we do this using the simplified flux factor in Eqn. (6.64).

$$d\sigma = \frac{m^4}{(2\pi)^2} \frac{\delta^4(p_{1'} - p_1 + p_{2'} - p_2)}{2E|\vec{p}|} \mid M_{fi} \mid^2_{SPIN} \frac{d^3 p_{1'}}{E_{1'}} \times 2\theta(p'_{20})\delta(p_2'^2 - m^2) \times d^4 p_{2'} \tag{6.69}$$

where we have cancelled some powers of (2π).

We now integrate $\delta^4(p_{1'} - p_1 + p_{2'} - p_2)$ over $d^4p_{2'}$. By definition of the Dirac δ function, this sets $p_{2'} = -p_{1'} + p_1 + p_2$ in the second Dirac δ function.

Using $p_{1'} = (E_{1'}, \vec{p}_{1'})$, $p_1 = (E, \vec{p})$ and $p_2 = (E, -\vec{p})$, we replace $p_{2'}$ by $(2E - E_{1'}, -\vec{p}_{1'})$ and rewrite the second Dirac δ function:

$$\delta(p_2'^2 - m^2) = \delta((p_1 + p_2 - p_{1'})^2 - m^2) \tag{6.70}$$

$$= \delta(4E^2 - 4EE_{1'} + E_1'^2 - |\vec{p}_1'|^2 - m^2) = \delta(4E^2 - 4EE_{1'})$$

Next we note that $\frac{d^3\vec{p}_{1'}}{E_{1'}}$ in Eqn 6.69 can be rewritten. Note that $E_1'^2 - m^2 = |\vec{p}_{1'}|^2$ means that $E_{1'}dE_{1'} = |\vec{p}_{1'}|d|\vec{p}_{1'}|$, which in turn means

$$d^3\vec{p}_{1'} = |\vec{p}_{1'}|^2 d|\vec{p}_{1'}|d\Omega_1 = |\vec{p}_{1'}|E_{1'}dE_{1'}d\Omega_{1'} \tag{6.71}$$

now using Eqn (6.71) in Eqn (6.69) to substitute for $d^3\vec{p}_{1'}$. Doing this we obtain:

$$d\sigma = \frac{m^4}{(2\pi)^2} \frac{\mid M_{fi} \mid^2_{SPIN}}{2E|\vec{p}|} \delta(4E^2 - 4EE_{1'}) 2|\vec{p}_{1'}|dE_{1'}d\Omega_{1'} \times \theta(2E - E_{1'}) \tag{6.72}$$

Note that we can now integrate over the δ function and $dE_{1'}$. Here we will pick up a factor of $\frac{1}{4E}$ since this is a δ function of a function of $E_{1'}$. Finally, for this stage, we can now use the kinematics discussed earlier to set $E_{1'} = E$ and $|\vec{p}_{1'}| = |\vec{p}|$. Doing this and inserting $\mid M_{fi} \mid^2_{SPIN}$, we get, after some cancellation, (and noting that the θ function is trivially satisfied) we get:

$$\frac{d\sigma}{d\Omega_{1'}} = \frac{e^4}{2E^2}\left\{\frac{(p_{1'}\cdot p_{2'})(p_1\cdot p_2)+(p_1\cdot p_{2'})(p_{1'}\cdot p_2)}{(p_1-p_{1'})^4}+\frac{2(p_{1'}\cdot p_{2'})(p_1\cdot p_2)}{(p_1-p_{1'})^2(p_{2'}-p_1)^2}\right. \tag{6.73}$$
$$\left.+\frac{(p_{2'}\cdot p_{1'})(p_1\cdot p_2)+(p_{2'}\cdot p_2)(p_1\cdot p_{1'})}{(p_{2'}-p_1)^4}\right\}$$

What remains to be done now is to compute all of the kinematical factors within the curly brackets. We note that $p_1 = (E, \vec{p})$, $p_2 = (E, -\vec{p})$, $p_{1'} = (E, +\vec{p}')$, $p_{2'} = (E, -\vec{p}')$ with $|\vec{p}| = |\vec{p}'|$ since $E, |\vec{p}| >> m$, we make the approximation (ultrarelativistic limit) that $E \approx |\vec{p}|$, of course as noted before.

The first term in the curly brackets is dealt with first. We begin with the numerator:

$$\begin{aligned} &(p_{1'}\cdot p_{2'})(p_1\cdot p_2)+(p_1\cdot p_{2'})(p_{1'}\cdot p_2) \\ &= (E^2+|\vec{p}'|^2)(E^2+|\vec{p}|^2)+(E^2+|\vec{p}||\vec{p}'|\cos\theta)(E^2+|\vec{p}||\vec{p}'|\cos\theta) \\ &= 4E^4+E^4(1+\cos\theta)^2 \end{aligned} \tag{6.74}$$

Since $|\vec{p}| \approx E$ and $|\vec{p}'| = |\vec{p}|$. Using the double angle formula, $\cos\theta = 2\cos^2\frac{\theta}{2} - 1$, Eqn 6.74 becomes

$$4E^4 + 4E^4\cos^4\frac{\theta}{2} \tag{6.75}$$

The denominator of the first term in curly brackets is $(p_1 - p_{1'})^4$ which is

$$\begin{aligned} &[(E-E,\vec{p}-\vec{p}')\cdot(E-E,\vec{p}-\vec{p}')]^2 = [(\vec{p}-\vec{p}')\cdot(\vec{p}-\vec{p}')]^2 \\ &= \left[|\vec{p}|^2+|\vec{p}'|^2-2|\vec{p}||\vec{p}'|\cos\theta\right]^2 \simeq (2E^2)^2(1-\cos\theta)^2 \end{aligned} \tag{6.76}$$

using $\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$ this becomes:

$$16E^4\sin^4\frac{\theta}{2} \tag{6.77}$$

The evaluation of the remaining scalar products and powers of the momentum transfers with the aforementioned approximations and use of the double angle formula is left to the reader. Combining everything into Eqn. 6.73, one obtains:

$$\frac{d\sigma}{d\Omega_{1'}} = \frac{e^4}{8E^2}\left\{\frac{1+\cos^4\frac{\theta}{2}}{\sin^4\frac{\theta}{2}}+\frac{2}{\sin^2\frac{\theta}{2}\cos^2\frac{\theta}{2}}+\frac{1+\sin^4\frac{\theta}{2}}{\cos^4\frac{\theta}{2}}\right\} \tag{6.78}$$

This is the differential cross section with respect to one electron's exit variable for a collision of two electrons in the center of mass frame.

6.5 Bhabha Scattering: $e^+e^- \to e^+e^-$

In order to consider $e^+e^- \to e^+e^-$ scattering we note that in chapter 4 we had developed a picture of positrons being equivalent electrons travelling backward in time, with a particular choice of contour using a representation of the propagator as a contour integral negative energy electrons travel backward in time whilst positive energy electrons forward in time hence the former depicted positrons and the latter electrons.

In the treatment of Moller ($e^-e^-e \to e^-e^-$) scattering we therefore considered only $E > 0$ electrons and hence all the scattering proceeded forward in time. The process of scattering forward in time from momentum and spin states labelled $1, 2$ to final states labelled $1', 2'$ is shown below in figure 6.2

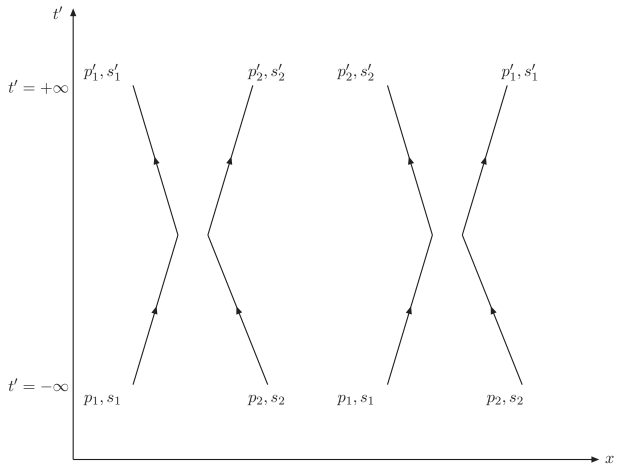

Fig. 6.2 Moller scattering: Two possible ways for a pair of indistinguishable electrons to scatter from initial momentum and spin states (p_1, s_1) and (p_2, s_2) to scatter forward in time into two final states of spin and momentum $(p_{1'}, s_{1'})$ and $(p_{2'}, s_{2'})$.

Now in order to consider $e^-e^+ \to e^-e^+$ we simply allow the electrons to travel backward as well as forward in time as discussed in previous chapters, the backward in time motion of an electron appears to us as a positron moving forwards in time. Let us associate label 2 with the negative energy solution, two things can happen now, the electron in the initial state labelled by 1 can scatter into the state labelled by $1'$ and the negative energy electron in the state $2'$ can scatter back in time into state 2, also the positive energy electron can start out propagating forward in time from state 1 and be *scattered backward in time to state 2* whilst the negative energy electron in state $2'$ initially propagating back in time can be *scattered by the potential forward in time into state* $1'$.

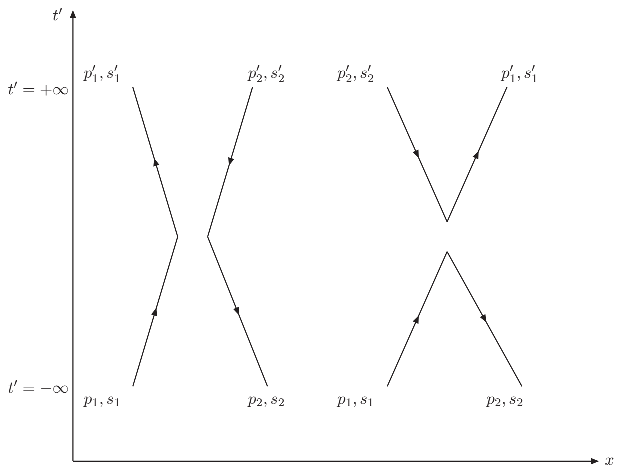

Fig. 6.3 A positive energy electron and a negative energy electron scattering. The states available to them are (p_1, s_1), $(\bar{p_2}, \bar{s_2})$ and $(p_{1'}, s_{1'})$ and $(\bar{p_{2'}}, \bar{s_{2'}})$. The label 2 signifies a negative energy state, these thus propagate back in time as depicted. The primed ($'$) states denote what we as observers see as the final state.

This process is now depicted in figure 6.3 which can be easily derived from figure 6.2

in which the amplitude on the left represents a change in momentum of one electron $p_{1'} - p_1$ and of the other by $\bar{p_2} - \bar{p_2}'$. In the process on the right the change in 4-momentum for one electron is $-\bar{p}_2' - p_1'$ and of the other by $p_1 + \bar{p}_2$ as is easily seen by considering the plane wave factors we would put in for the negative energy solutions as was done in section 6.1 and following the discussion thereafter. It is easy to see that this description depicted in figure 6.3 when translated into the language of positrons describes a electron positron pair annihilating "into" a virtual photon-and then another pair reappearing-this done by reversing the directions of the arrows on the legs corresponding to backward in time propagation.

The observer with a normal sense of time will see an electron and positron pair in initial spin momentum states $(p_1, s_1,)$ and $(\bar{p}_2, \bar{s}_2)$ scatter forward in time into an electron and positron state of spin and momenta $(p_{1'}, s_{1'})$ and $(\bar{p_{2'}}, \bar{s_{2'}})$.

By looking at the change of momenta we discussed and comparing these to Moller scattering we can see that the signs of the momenta of the positrons have been reversed. This is exactly what we would expect from the backward in time propagation. (CAREFUL!!!!-¿ space time, better perhaps to bring in the 4-velocity and write this down ?).

We can now transform the matrix element in 6.24 to the appropriate one for electron positron scattering by comparing the two processes in Fig. 6.3 and making appropriate modifications. Using 2 to denote the negative energy solutions, the initial states are thus labelled by $\bar{2}'$ and the final states by $\bar{2}$, and so using the appropriate negative energy solutions we will obtain:

For the sake of convention adopted in several texts we make the change $(p_2, s_2) \rightarrow (\bar{p_2}, \bar{s_2})$ and $(p_{2'}, s_{2'}) \rightarrow (\bar{p_2}', \bar{s_2}')$. Now, we've described the scattering process in the language of electrons.

Chapter 7

Bremsstrahlung

We will now calculate the cross-section for an electron scattering in the field of a nucleus to emit a photon. This process is called "Bremsstrahlung" or "braking" radiation (German). A little thought should make it clear to the reader that we will require the 4-vector potential of a free photon as well as that of the nucleus (we have covered the coulomb potential of a nucleus in Chapter 5). We note that the 4-vector potential of a freely propagating photon observed in a region of space far away from charges and currents will satisfy:

$$\Box A^{\mu}(x) = 0 \tag{7.1}$$

This leads to the solution:

$$A^{\mu}(x) = N\epsilon^{\mu}(e^{ik\cdot x} + e^{-ik\cdot x}) \tag{7.2}$$

We note the following regarding Eqn. 7.2:

1. Here, k is the photon's 4-momentum satsifying:

   $$k^2 = (k^0)^2 - \vec{k}\cdot\vec{k} = \omega^2 - \vec{k}\cdot\vec{k} = 0$$

   since the photon is massless. This in turn leads to the result that the photon energy $\omega = |\vec{k}|$. The direction of $\vec{k}$ is, of course, the direction of propagation of the photon.

2. The unit vector ϵ^{μ} is a four-vector; we shall simplify and modify this shortly.

3. Finally, N in Eqn. 7.2 is a normalization of the vector potential so that it carries the appropriate energy in the volume V specified by the time averaged Poynting vector. We will determine this as well.

Since we have freedom in the choice of gauge in describing $A^\mu(x)$ (see discussion preceding and including Eqn. 1.77), we can simply require $\partial_\mu A^\mu(x) = 0$ and obtain:

$$\begin{aligned} &k_\mu \epsilon^\mu N(e^{ik\cdot x} + e^{ik\cdot x}) = 0 \\ &k_\mu \epsilon^\mu = k \cdot \epsilon = 0 \end{aligned} \tag{7.3}$$

Note now that since $k \cdot k = 0$, we can always add a 4-vector proportional to k^μ to ϵ^μ and still preserve the gauge condition since :

$$k_\mu(\epsilon^\mu + \lambda k^\mu) = \quad k \cdot \epsilon + \lambda k^2 = k \cdot \epsilon = 0 \tag{7.4}$$

We can choose a value of the constant parameter λ so that the time component of ϵ^μ is exactly zero and then ϵ is purely space like and so we can say:

$$\begin{aligned} &\vec{\epsilon} \cdot \vec{k} = 0 \text{ or} \\ &\vec{\nabla} \cdot \vec{A} = 0 \end{aligned} \tag{7.5}$$

This is the radiation gauge condition, and since $\vec{k}$ is the direction of the photon's propagation, it means that the polarization vector ϵ is perpendicular to it and lies in a two-dimensional plane and so the photon has two linearly independent polarizations.

Onward to the normalization factor N. Recall that the energy in an electro-magnetic wave is $\frac{1}{8\pi}$ times the integral over space of the time averaged energy density given by:

$$\frac{1}{8\pi} < E^2 + B^2 >_{\text{time}}$$

In keeping with our convention that all observed particles are in a box of volume V, we can now insert the previously mentioned integral over space:

$$\frac{1}{8\pi} \int_V < E^2 + B^2 >_{\text{time}} d^3x$$

Now we turn to the vector potential $A^\mu(x)$ to calculate B :

$$A^\mu(x, k) = \epsilon^\mu N \left\{ e^{-ik\cdot x} + e^{ik\cdot x} \right\}$$

$$\vec{B} = \vec{\nabla} \times \vec{A} = iN\vec{k} \times \vec{\epsilon} \left[e^{-ik\cdot x} - e^{ik\cdot x} \right]$$

$$= 2N(\vec{k} \times \vec{\epsilon}) \sin kx$$

$$\therefore B^2 = 4N^2(\vec{k} \times \vec{\epsilon}) \cdot (\vec{k} \times \vec{\epsilon}) \sin^2 \vec{k} \cdot \vec{x}$$

.

The cross product is easily done and the scalar product yields:

$$(\vec{k} \times \vec{\epsilon}) \cdot (\vec{k} \times \vec{\epsilon}) = (\vec{k} \cdot \vec{k})(\vec{\epsilon} \cdot \vec{\epsilon}) - (\vec{k} \cdot \vec{\epsilon})(\vec{k} \cdot \vec{\epsilon}) = |\vec{k}|^2$$

using the radiation gauge condition 7.5 this gives

$$|B|^2 = 4N^2|\vec{k}|^2 \sin^2(\vec{k} \cdot \vec{x} - \omega t)$$

note that similarly we can use

$$E = -\vec{\nabla}\phi(x,t) - \frac{1}{c}\frac{\partial \vec{A}((x,t)}{\partial t}$$

where $\phi(x,t) = 0$ in this region of space and this gives us

$$|E|^2 = 4N^2\frac{\omega^2}{c^2} \sin^2(\vec{k} \cdot \vec{x} - \omega t)$$

of course in our system of units $c = 1$ and so we can add up the contributions from from the E and B fields to the energy density and so the energy of the photon can be written as:

$$E_{\text{photon}} = \frac{4\omega^2 N^2}{4\pi} \int_V d^3x < \sin^2(\omega t - \vec{k} \cdot \vec{x}) >_{\text{time}}$$

the integral over space introduces a factor of V and we know the time average of $\sin^2(\vec{k} \cdot x - \omega t)$ is simply $\frac{1}{2}$ this means that the energy of a photon is given by:

$$E_{\text{photon}} = \frac{\omega^2 N^2}{2\pi} \times V,$$

now

$$E_{\text{photon}} = \omega$$

and so

$$\omega = \frac{\omega^2 N^2}{2\pi} V$$

which in turn means that:

$$N = \sqrt{\frac{2\pi}{\omega V}}$$

and so finally we can write the vector potential representing the photon as:

$$A^\mu(x) = \left(e^{ik\cdot x} + e^{-ik\cdot x}\right)\sqrt{\frac{2\pi}{\omega V}}\epsilon^\mu \tag{7.6}$$

Lets write down the potential of the nucleus and of the emitted photon. The Coulomb potential of our nucleus is:

$$A_Z^0(x) = -\frac{Ze}{|\vec{x}|} \tag{7.7}$$

and the potential of the radiated photon is:

$$A_\gamma^\mu(x) = \left(\frac{2\pi}{\omega V}\right)^{\frac{1}{2}}(e^{ik\cdot x} + e^{-ik\cdot x})\epsilon^\mu \tag{7.8}$$

where the 4-momentum is k with $k^2 = 0$.

Note the use of the subscripts Z and γ to denote the Coulombic and photon potentials. The scattering amplitude series into which we will insert the sum of A_Z^0 and A_γ^μ is as always:

$$\begin{aligned} S_{fi} = \delta_{fi} - ie\int d^4x_1\bar{\psi}_f(x_1)\not{A}(x_1)\psi_i(x_1) \\ - ie^2\int d^4x_1\int d^4x_2\bar{\psi}_f(x_2)\not{A}(x_2)S_F(x_2-x_1)\not{A}(x_1)\psi_i(x_1) + \cdots \end{aligned} \tag{7.9}$$

where $\not{A}(x_{1,2}) = \gamma^0 A_{0,Z}(x_{1,2}) + \gamma^\mu A_{\mu,\gamma}(x_{1,2})$ to first order we obtain

$$-ie\int d^4x_1\left\{\bar{\psi}_f(x_1)\left[\gamma_0 A_Z^0(x_1) + \gamma_\mu A_\gamma^\mu(x_1)\right]\psi_i(x_1)\right\}$$

$$= -ie\int d^4x_1\bar{\psi}_f(x_1)\gamma^0 A_Z^0(x_1)\psi_i(x_1) - ie\int d^4x_1\bar{\psi}_f(x_1)\gamma_\mu A_\gamma^\mu(x_1)\psi_i(x_1)$$

This first piece of the amplitude is exactly what we obtained in our calculation of Mott scattering.

Let's examine the second amplitude a little more closely

$$-ie\int d^4x_1\bar{\psi}_f(x_1)\gamma_\mu A_\gamma^\mu(x_1)\psi_i(x_1)$$

$$-ie\int d^4x_1\psi_f(x_1)\gamma_\mu\sqrt{\frac{2\pi}{\omega V}}\epsilon^\mu[e^{ik\cdot x_1} + e^{-ik\cdot x_1}]\psi_i(x_1)$$

$$-ie\int d^4x_1\sqrt{\frac{m}{E_fV}}\bar{u}_f e^{ip_f\cdot x_1}\gamma_\mu\epsilon^\mu\sqrt{\frac{2\pi}{\omega V}}[e^{ik\cdot x_1}+e^{-ik\cdot x_1}]\times\sqrt{\frac{m}{E_iV}}\times e^{-ip_i\cdot x_1}u_i$$

where the explicit plane factors and spinors have been inserted for the ψs

Let's just examine one of the pieces of the above:

$$-ie\int d^4x_1\sqrt{\frac{m^2}{E_fE_iV^2}}\sqrt{\frac{2\pi}{\omega V}}e^{ip_f\cdot x_1+ik\cdot x_1-ip_i\cdot x_1}\bar{u}_f\gamma_\mu\epsilon^\mu u_i$$

the integral of the exponential term over the variable x_1 forms a 4-dimensional Dirac delta function: $(2\pi)^4\delta^4(p_f+k-p_i)$. This states that the difference between the initial and final electron momenta $= -k$, the 4-momentum of the photon.

$$p_f - p_i = -k$$

Taking the scalar product of both sides with respect to themselves we obtain

$$(p_f - p_i)\cdot(p_f - p_i) = k^2 \tag{7.10}$$

the photon mass is 0 and the electron mass is m :

$$p_f^2 + p_i^2 - 2p_f\cdot p_i = 0$$

$$2m^2 - 2(E_fE_i - \vec{p}_f\cdot\vec{p}_i) = 0$$

This can be viewed in any Lorentz frame. To simplify matters, let us view this in a frame where the electron comes to rest ($\vec{p}_f = 0$, $E_f = m$). Thus

$$k^2 = 2m^2 - 2(E_fE_i - \vec{p}_f\cdot\vec{p}_i) = 2m^2 - 2m(E_i) \tag{7.11}$$

It is easy to see that the above equation can only be satisfied for $k^2 = 0$ (since free photons are massless) if $E_i = m$ or if the electron is at rest in its initial and final states. From 4-momentum conservation, this would then mean that the radiated photon has energy and spatial momentum all equal to zero. It's clear that we cannot demand that a photon is emitted and equate the change of momentum of the electron to the photon. The two processes that arise out of the first order term represent the sum of the amplitudes of Mott scattering and the physically impossible emission of a photon from an electron that changes its momentum. Neither of these amplitudes describe the physical process whose cross section we are trying to calculate and one of these amplitudes carries an energy momentum conservation condition that cannot be satisfied.

In order for the condition $k^2 = 0$ to be satisfied in Eqn. 7.11 we need another quantity that allows the initial electron to be in motion. We interpret this as a 4-momentum imparted by another agent, this we assume to be the field of the nucleus. Also the form of the amplitudes do not represent both a scattering in a Coulomb potential and the emission of a photon. We reject the first order terms and go to second order. At second order, as we shall see below, the field of the nucleus is the agent imparting the momentum to the electron which allows the photon to be emitted. Also the second order term contains two terms that can easily be interpreted as physically relevant

$$-ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2) \not{A}(x_2) S_F(x_2 - x_1) \not{A}(x_1) \psi_i(x_1) \tag{7.12}$$

We now replace $A(x_{1,2})$ in Eqn. 7.12 with $A^0_Z(x_{1,2}) + A^\mu_\gamma(x_{1,2})$ to obtain:

$$-ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2)[A^0_Z(x_2)\gamma_0 + A^\mu_\gamma(x_2)\gamma_\mu] S_F(x_2 - x_1)[A^0_Z(x_1)\gamma_0 + A^\mu_\gamma(x_1)\gamma_\mu]\psi_i(x_1) \tag{7.13}$$

Note that Eqn. 7.13 will give rise to terms with two free photon fields. These can easily be identified as Mott scattering to second order. We select only the amplitudes that contain one Coulombic and one free photon potential as being physically relevant:

$$\begin{aligned} &- ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2) A^0_Z(x_2)\gamma_0 S_F(x_2 - x_1) A^\mu_\gamma(x_1)\gamma_\mu \psi_i(x_1) \\ &- ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2) A^\mu_\gamma(x_2)\gamma_\mu S_F(x_2 - x_1)\gamma_0 A^0_Z(x_1)\psi_i(x_1) \end{aligned} \tag{7.14}$$

Using the explicit form of the potentials and spinors ($\psi_l(x) = e^{-ip_l \cdot x}\sqrt{\frac{m}{E_l V}} u_{l,s}$ as in previous chapters), we obtain after collecting some terms:

$$-ie^2 \int \frac{d^4p}{(2\pi)^4} \int d^4x_1 \int d^4x_2 \sqrt{\frac{m^2}{E_f E_i}} \frac{1}{V} \sqrt{\frac{2\pi}{\omega V}} (-Ze)$$

$$\times \bar{u}_f e^{ip_f \cdot x_2} \left\{ \frac{\gamma^0}{|\vec{x}_2|} \times \frac{(\not{p} + m) e^{-ip \cdot (x_2 - x_1)}}{p^2 - m^2 + i\epsilon} \cdot \gamma^\mu \epsilon_\mu \times [e^{ik \cdot x_1} + e^{-ik \cdot x_1}] \right.$$

$$+e^{-ip\cdot(x_2-x_1)}\gamma^\mu\epsilon_\mu[e^{ik\cdot x_2}+e^{-ik\cdot x_2}]\frac{\not p+m}{p^2-m^2+i\epsilon}\frac{\gamma^0}{|\vec{x}_1|}\Bigg\}e^{-ip_i\cdot x_1}u_i$$

where we have used the expression for the Feynman propagator:

$$S_F(x_2-x_1)=\int\frac{d^4p}{(2\pi)^4}e^{-ip\cdot(x_2-x_1)}\cdot\frac{\not p+m}{p^2-m^2+i\epsilon}$$

It should be clear that the amplitude is the some of two processes: one where the Coulomb potential acts first and then the photon is radiated, and the second where the photon is radiated before the Coulomb potential acts.

We simplify the factors a little further to obtain:

$$ie^3Z\sqrt{\frac{m^2}{E_fE_i}}\times\frac{1}{V^{\frac{3}{2}}}\times\sqrt{\frac{2\pi}{\omega}}\int d^4x_1\int d^4x_2\int\frac{d^4p}{(2\pi)^4}\times \tag{7.15}$$

$$\bar{u}_fe^{ip_f\cdot x_2}\Bigg\{\frac{\gamma^0}{|\vec{x}_2|}\cdot\frac{\not p+m}{p^2-m^2+i\epsilon}e^{-ip\cdot(x_2-x_1)}\gamma^\mu\epsilon_\mu\times(e^{ik\cdot x_1}+e^{-ik\cdot x_1})$$

$$+\gamma^\mu\epsilon_\mu[e^{ik\cdot x_2}+e^{-ik\cdot x_2}]\frac{\not p+m}{p^2-m^2+i\epsilon}\cdot e^{-ip\cdot(x_2-x_1)}\cdot\frac{\gamma^0}{|\vec{x}_1|}\Bigg\}e^{-ip_i\cdot x_1}u_i$$

The reader should by now immediately recognize that the integrals over either dx_1 or dx_2 should produce Dirac δ functions when taken over the appropriate exponential terms. Let us begin by considering the first term of the expansion in Eq. 7.15: Let's start by collecting all of the exponents and integrals for *one* of the two terms-the one corresponding to $e^{+ik\cdot x_1}$ in the sum

$$[e^{ik\cdot x_1}+e^{-ik\cdot x_1}]$$

. This choice will be explained later.

Collecting only the terms in Eqn. 7.15 that have integrals performed on in d^4x_1, d^4x_2, d^4p and keeping terms with gamma matrices in them to keep the ordering:

$$\int\frac{d^4p}{(2\pi)^4}\int d^4x_1\int d^4x_2e^{ip_f\cdot x_2-ip\cdot x_2+ip\cdot x_1+ik\cdot x_1-ip_i\cdot x_1}\times\frac{1}{|\vec{x}_2|}\times\gamma^0\frac{\not p+m}{p^2-m^2+i\epsilon}\gamma^\mu\epsilon_\mu$$

This gives us, after separating the x_2 and x_1 terms:

$$\int \frac{d^4p}{(2\pi)^4} \int d^4x_1 \int d^4x_2 e^{ip_f\cdot x_2 - ip_i\cdot x_2} \frac{1}{|\vec{x}_2|} \times e^{ip\cdot x_1 + ik\cdot x_1 - ip_i\cdot x_1} \gamma^0 \times \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} \gamma^\mu \epsilon_\mu$$

We can't form a δ function from the integral over d^4x_2 due to the presence of the $\frac{1}{|\vec{x}_2|}$ factor but can do this over the d^4x_1 (a similar argument applies for $2 \leftrightarrow 1$):

$$\int \frac{d^4p}{(2\pi)^4} \int dx_2 e^{ip_f\cdot x_2 - ip\cdot x_2} \frac{1}{|\vec{x}_2|} \times (2\pi)^4 \delta^4(p + k - p_i) \times \gamma^0 \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} \gamma^\mu \epsilon_\mu$$

One can now integrate over d^4p and set $p = p_i - k$ everywhere in the above expression. Doing this we obtain:

$$\int d^4x_2 e^{ip_f\cdot x_2 - ip\cdot x_2 + ik\cdot x_2} \frac{1}{|\vec{x}_2|} \gamma^0 \times \frac{\not{p}_i - \not{k} + m}{(p_i - k)^2 - m^2 + i\epsilon} \gamma^\mu \epsilon_\mu \qquad (7.16)$$

By breaking up the integral over d^4x_2 into the space and time pieces and identifying $E_f = p_{f0}$, $E_i = p_{i0}$ and $\omega = k_0$ we obtain for the above:

$$2\pi\delta(E_f + \omega - E_i) \int d^3x_2 \frac{e^{-i(\vec{p}_f - \vec{p}_i + \vec{k})\cdot \vec{x}_2}}{|\vec{x}_2|} \times \frac{\not{p}_i - \not{k} + m}{(p_i - k)^2 - m^2 + i\epsilon} \gamma^\mu \epsilon_\mu \quad (7.17)$$

Why did we ignore the $e^{-ik\cdot x_1}$ term in $[e^{ik\cdot x_1} + e^{-ik\cdot x_1}]$ when gathering exponentials in Eqn. 7.15? The reader can verify that $e^{ik\cdot x_1}$ has led us to the Dirac δ-function in energy in Eqn. 7.17, i.e. $2\pi\delta(E_f - E_i + \omega)$. This δ function is consistent with our requirement that photon emission take place, i.e. that the final energy is less than initial energy: $E_f = E_i - \omega$ which is the requirement imposed by the Dirac δ function. We will heretofore keep $e^{+ik\cdot x}$ for emitted and $e^{-ik\cdot x}$ for absorbed photons.

Back to Eqn. 7.17. Let's look at the integral over d^3x_2 :

$$\int d^3x_2 \frac{e^{-i(\vec{p}_f - \vec{p}_i + \vec{k})\cdot x_2}}{|\vec{x}_2|}$$

We have encountered this before in our calculation of Mott scattering in the discussion leading up to Chapter 5 Eqn. 5.14. Setting $|\vec{q}| = |\vec{p}_f - \vec{p}_i + \vec{k}|$ we have:

$$\int d^3x_2 \frac{e^{-i\vec{q}\cdot\vec{x}}}{|\vec{x}_2|} = \frac{4\pi}{|\vec{q}|^2}$$

It is quite straightforward to complete this procedure for the second term in Eqn. 7.15 to obtain:

$$ie^3 Z\sqrt{\frac{m^2}{E_f E_i}}\frac{1}{V^{\frac{3}{2}}}\sqrt{\frac{2\pi}{\omega}}\frac{4\pi}{|\vec{q}|^2}\bar{u}_f\left[\frac{\gamma^0(\not{p}_i - \not{k} + m)\not{\epsilon}}{(p_i-k)^2-m^2+i\epsilon}+\frac{\not{\epsilon}(\not{p}_f+\not{k}+m)\gamma^0}{(p_f+k)^2-m^2+i\epsilon}\right]u_i \tag{7.18}$$

$$\times 2\pi\delta(E_f - E_i + \omega)$$

This can be rewritten as:

$$\sqrt{\frac{m^2}{E_f E_i}}\sqrt{\frac{2\pi}{\omega}}\frac{1}{V^{\frac{3}{2}}}M_{fi}2\pi\delta(E_f+\omega-E_i) \tag{7.19}$$

where

$$M_{fi}=\frac{iZe^3 4\pi}{|\vec{q}|^2}\times\bar{u}_f\left[\frac{\gamma^0(\not{p}_i - \not{k} + m)\not{\epsilon}}{(p_i-k)^2-m^2+i\epsilon}+\frac{\not{\epsilon}(\not{p}_f+\not{k}+m)\gamma^0}{(p_f+k)^2-m^2+i\epsilon}\right]u_i \tag{7.20}$$

LEFTOVER: Use above expressions for S_{fi} and M_{fi} to justify Feynmann rules. Reintroduce (-i) factors etc. Tie in with Mott scattering Coulomb potential. Also this is the expression you (probably or something similar) or something similar you want to pick up for pair production. (Before soft photon limit).

The Eqn (7.20) can be simplified before we proceed to square the transition amplitude and insert all counting factors and fluxes. These simplifications are as follows:

1. We first look at the denominators of the two terms within square brackets in Eqn. 7.20. $(p_i-k)^2-m^2 = -2p_i\cdot k+k^2+m^2-m^2$. Since $k^2=0$, this is simply $-2p_i\cdot k$. Similarly $(p_f+k)^2-m^2 = 2p_f\cdot k$.
2. Next we state that we will calculate Bremsstrahlung in the limit $k\approx 0$ (soft photon limit), we thus set all k in the *numerator* of Eqn 7.20 to zero. Next we simplify the numerators of the terms within square brackets of Eqn. 7.20-reminding ourselves that $k\approx 0$ in the numerators:
   i. The numerator of the first term in the square brackets of Eqn. 7.20 becomes: $\not{p}_i\ \not{\epsilon}$ which can be re-written using the anticommutation relations of the gamma matrices:
   $$\gamma^\mu\gamma^\nu p_{i\mu}\epsilon_\nu = (2g^{\mu\nu}-\gamma^\nu\gamma^\mu)p_{i\mu}\epsilon_\nu = 2p_i\cdot\epsilon - \not{\epsilon}\not{p}_i$$
   ii. A similar manipulation of the second numerator yields:
   $$\not{\epsilon}\not{p}_f = 2p_f\cdot\epsilon - \not{p}_f\ \not{\epsilon}.$$

With these modifications, the term within square brackets in Eqn. 7.20 can now be rewritten :

$$\frac{\bar{u}_f \not\epsilon(\not p_i + m)u_i + 2p_i \cdot \epsilon \bar{u}_f \gamma^0 u_i}{-2p_i \cdot k} + \frac{\bar{u}_f(-\not p_f + m)\not\epsilon u_i + \bar{u}_f \gamma^0 u_i \cdot 2p_f \cdot \epsilon}{2p_f \cdot k} \tag{7.21}$$

Note that the adjoint spinor satsifies the Dirac equation $\bar{u}(\not p - m) = 0$. A spinor satisfies the Dirac equation $(\not p - m)u = 0$. Using these in we see that Eqn. 7.21 $(-\not p_i + m)u_i$ and $\bar{u}_f(-\not p_f + m)$ both equal zero we now rewrite M_{fi} :

$$M_{fi} = \frac{Ze^3 4\pi}{|\vec{q}|^2} \times \left[\frac{p_f \cdot \epsilon}{p_f \cdot k} - \frac{p_i \cdot \epsilon}{p_i \cdot k}\right] \times \bar{u}_f \gamma^0 u_i \tag{7.22}$$

The reader should at this point note that we will have to square this expression along with the remaining factors in Eqn (7.19) and insert the appropriate flux and counting factors to obtain a differential cross section. For this, the modified M_{fi} in Eqn. 7.22 is inserted into Eqn. 7.19 and the whole expression is multiplied by its complex conjugate. To get the differential cross section from this, the following factors are inserted–with which the reader should by now be familiar :

1. The counting factors for the outgoing electron and photon which are are respectively: $\frac{Vd^3\vec{p}_f}{(2\pi)^3}$ and $\frac{Vd^3\vec{k}}{(2\pi)^3}$. These can be simplified further by using $d^3|\vec{p}| = |\vec{p}|^2 d|\vec{p}| d\Omega = |\vec{p}| E dE d\Omega$ and remembering that for the photon $|\vec{k}| = \omega$ (see discussion just preceding Eqn.5.26 upto and including Eqn. 5.28):
   i. The counting factor for the exiting electron $\frac{Vd^3\vec{p}_f}{(2\pi)^3}$ can then be replaced by $\frac{V|\vec{p}_f|E_f dE_f d\Omega_f}{(2\pi)^3}$
   ii. The counting factor for the exiting photon $\frac{Vd^3\vec{k}}{(2\pi)^3}$ can then be replaced by $\frac{V\omega^2 d\omega d\Omega_k}{(2\pi)^3}$
2. The square of 2π times a Dirac delta function in energy (see discussion preceding Eqn. 5.22) is replaced by $(2\pi)^2 \frac{T}{2\pi}\delta(E_f - E_i)$; both sides of the resulting equation are then divided by T which yields a differential element of rate $dR = |\vec{j}_i| \, d\sigma$.
3. The flux factor for incoming electrons was already derived in the Mott scattering chapter (discussion preceding and including Eqn. 5.33) and is $\frac{VE_i}{|\vec{p}_i|}$. We must divide by this to turn the rate into a differential cross section.

4. Since we are using the the soft photon limit, $|\vec{p}_f| \approx |\vec{p}_i|$.

After the insertion of all these factors we obtain:

$$d\sigma = 4e^4 Z^2 m^2 \frac{|\bar{u}_f \gamma^0 u_i|^2}{|\vec{q}|^4} \left[\frac{p_f \cdot \epsilon}{p_f \cdot k} - \frac{p_i \cdot \epsilon}{p_i \cdot k}\right]^2 \frac{\omega e^2}{(2\pi)^2} \times \delta(E_f + \omega - E_i) \quad (7.23)$$
$$\times dE_f d\Omega_f d\omega d\Omega_k$$

Identifying the Mott cross section from our previous work (Eqn. 5.35) we can write after integration over the δ function:

$$\frac{d\sigma}{d\Omega_k d\Omega_f d\omega} = \theta(E_f + \omega - E_i) \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{MOTT}} \times \left[\frac{p_f \cdot \epsilon}{p_f \cdot k} - \frac{p_i \cdot \epsilon}{p_i \cdot k}\right]^2 \frac{\omega e^2}{(2\pi)^2} \quad (7.24)$$

LEFTOVER: Replace MOTT with elastic ? The θ function expresses the fact that the initial energy must be greater than the final by at least ω-the photon energy.

Note that it is possible to consider a case where we integrate over all final photon energies, which means integrating (7.24) with respect to $d\omega$. The numerator in Eqn. 7.24 contains a term proportional to ω and the denominator, by virtue of $|\vec{k}|^2 = \omega^2$ a term of ω^2. This resulting integral of $\frac{d\omega}{\omega}$ is divergent near $\omega = 0$. This is the so called "infrared catastrophe". It is interesting to examine its origins. We begin this calculation by going to second order in a process where we demand that a photon is emitted but let its energy go to zero. However, it is also possible to calculate Mott scattering to second order, where no photon is emitted. It turns out that if this is done, the divergent piece discussed above is exactly cancelled by the divergent piece of the second order Mott calculation. This second order process *must* contain the interaction of the electron with its own field. This is discussed in a later chapter, but is at the boundaries of the scope of this book. We now limit the range of integration of Eqn. 7.24 and present the final result of Bremsstrahlung.

Since we do not, in practise, observe the final polarization states of the photon we must sum over these. In order to do this, we first isolate the relevant piece from Eqn 7.24: $|\frac{p_f \cdot \epsilon}{p_f \cdot k} - \frac{p_i \cdot \epsilon}{p_i \cdot k}|^2$. The operation we must perform is

$$\sum_{\lambda=1}^{2} \epsilon_\mu(\lambda, k) \epsilon_\nu^*(\lambda, k) M^\mu M^\nu \quad (7.25)$$

where we have used

$$M^\mu = \frac{p_f^\mu}{p_f \cdot k} - \frac{p_i^\mu}{p_i \cdot k} \tag{7.26}$$

and $\epsilon^\mu(\lambda, k)$ represents the μ^{th} component of the polarization λ (each polarization state is labelled by a λ) of the photon which carries 4-momentum k.

We can simplify the sum in (7.25) by noting that

$$k_\mu M^\mu = k \cdot M = 0 \tag{7.27}$$

Proceeding further with the simplification, we pick one axis (x for example) to lie along $\vec{k}$ and we obtain, using (7.27), $kM^0 - kM^1 = 0$. This means that Eqn. 7.27, implies that $M^0 = M^1$. Of course picking $\vec{k}$ to point along along x, the polarization vectors of the photon must lie in the $y - z$ plane. We pick $\epsilon_\mu(\lambda = 1, 2, k)$ to lie along y and z. with this choice we find that Eqn. 7.25 becomes trivially:

$$M^2M^2 + M^3M^3 \tag{7.28}$$

Note that since (7.27) implies $M^0 = M^1$ (7.28) can be written as

$$M^2M^2 + M^3M^3 + M^1M^1 - M^0M^0 = -g_{\mu\nu}M^\mu M^\nu$$

This leads us to conclude that $\sum_{\lambda=1}^{2} \epsilon^\mu(\lambda)\epsilon^{\nu*}(\lambda) = -g^{\mu\nu}$ for use in expressions such as $\sum \epsilon^\mu(x)\epsilon^{\nu*}(x)M_\mu M_\nu$ *if* $k_\mu M^\mu = 0$ *and* because of our choice of gauge. Now

$$-g^{\mu\nu}M_\mu M_\nu = -\left(\frac{p_f}{k \cdot p_f} - \frac{p_i}{k \cdot p_i}\right)_\mu \left(\frac{p_f}{k \cdot p_f} - \frac{p_i}{k \cdot p_i}\right)^\mu$$

and this gives us:

$$\frac{2p_f \cdot p_i}{(k \cdot p_f)(k \cdot p_i)} - \frac{m^2}{(k \cdot p_f)^2} - \frac{m^2}{(k \cdot p_i)^2} \tag{7.29}$$

Using $p_{f,i} = (E_{f,i}, \vec{p}_{f,i}) = (m\gamma_{f,i}, m\vec{\beta}_{f,i}, \gamma_{f,i})$ and $k = (\omega, \vec{k})$ with $|\vec{k}| = \omega$ and $\hat{k} - \frac{\vec{k}}{|\vec{k}|}$ we obtain the following simplified expression:

$$\frac{2(1 - \vec{\beta}_f \cdot \vec{\beta}_i)}{\omega^2(1 - \hat{k} \cdot \vec{\beta}_f)(1 - \hat{k}\vec{\beta}_i)} - \frac{m^2}{\omega^2 E_f^2(1 - \hat{k} \cdot \vec{\beta}_f)^2} - \frac{m^2}{\omega^2(1 - \hat{k} \cdot \vec{\beta}_i)E_i^2} \tag{7.30}$$

We have to now insert this back into equation (7.24) and integrate this over the final photon exit solid angle and energy. However we have to take care to restrict the integration over a range of ω to avoid the "infrared catastrophe" stemming from our neglect of the second order contribution from the elastic process. We note that there are several scalar products in Eqn. 7.30. We are trying to integrate over $d\Omega_k$ and $d\omega$, and obtain $\frac{d\sigma}{d\Omega_f}$ which is in terms of the electron's final momentum direction. We define θ to be the angle between $\vec{\beta}_f$ and $\vec{\beta}_i$ and take the soft photon approximation, thus $|\vec{\beta}_f| \approx |\vec{\beta}_i| = |\vec{\beta}|$ and so $1 - \vec{\beta}_f \cdot \vec{\beta}_i = 1 - \beta^2 \cos\theta$.

Next, regarding the terms containing a scalar product of $\hat{k}$ with one of the $\vec{\beta}$s, we note that we can define the z axis for the definition of $d\Omega_k$ to lie along any arbitrary direction: specifically any combinations of the $\vec{\beta}$s. Since we are integrating over all of $d\Omega_k$, this doesn't matter. We will combine all of Eqn. 7.30 into Eqn 7.24. Note that a factor of w will cancel leaving one in the denominator; note also the soft photon approximation has been applied, thus $E_f \approx E_i$, $|\vec{\beta}_f| \approx |\vec{\beta}_i| = |\vec{\beta}|$. Doing all this and re-arranging, we obtain:

$$\frac{d\sigma}{d\Omega_f} = \left(\frac{d\sigma}{d\Omega_f}\right)_{\text{ELASTIC}} \times \frac{e^2}{(2\pi)^2} \times \ln\left(\frac{\omega_{\text{MIN}}}{\omega_{\text{MAX}}}\right) \tag{7.31}$$

$$\times \left[\frac{2(1-\beta^2\cos\theta)}{(1-\hat{k}\cdot\vec{\beta}_f)(1-\hat{k}\cdot\vec{\beta}_i)} - \frac{m^2}{E^2(1-\hat{k}\cdot\vec{\beta}_f)^2} - \frac{m^2}{E^2(1-\hat{k}\cdot\vec{\beta}_i)^2}\right] \times d\Omega_k$$

It should be obvious to the reader that the integration over both the m^2 terms will yield one result. With χ defined as the angle between $\hat{k}$ and $\vec{\beta}_f$ we obtain and $\vec{\beta}_f$ pointing along z:

$$\int d\Omega_k \frac{m^2}{E^2(1-\cos\chi\beta)^2} = 2\pi\frac{m^2}{E^2}\int_{-1}^{1}\frac{d(\cos\chi)}{(1-\beta\cos\chi)^2}$$

Changing limits with $u = 1 - \beta\cos\chi$, we obtain:

$$\frac{(-)}{\beta}2\pi\frac{m^2}{E^2}\int_{1+\beta}^{1-\beta}\frac{du}{u^2} = \frac{-2\pi}{-\beta}\frac{m^2}{E^2} \times \frac{2\beta}{(1-\beta^2)} = 4\pi$$

the second m^2 term will be the same, so we now factor out the $2(1-\beta^2\cos\theta)$ and concentrate on:

$$\int d\Omega_k \times \frac{1}{(1-\hat{k}\cdot\vec{\beta}_f)(1-\hat{k}\cdot\vec{\beta}_i)}$$

Here we will use a standard trick due to Feynman and begin by expressing this integral using

$$\frac{1}{ab} = \int_0^1 \frac{dx}{[ax + (1-x)b]^2}$$

with $a = 1 - \hat{k} \cdot \vec{\beta}_f$ and $b = 1 - \hat{k} \cdot \vec{\beta}_i$.

We must consider therefore:

$$\int_0^1 \frac{dx d\Omega_k}{[(1 - \hat{k} \cdot \vec{\beta}_f)x + (1 - \hat{k} \cdot \vec{\beta}_i)(1-x)]^2}$$

As in the integration of the m^2 terms over $d\Omega_k$, we redefine the solid angle and hence $\cos\chi$ to be the angle between $x\vec{\beta}_f + (1-x)\vec{\beta}_i$ and $\hat{k}$ which we take to be the direction of $\hat{z}$ for this integration. In an exactly analogous fashion we obtain:

$$2\pi \int_0^1 dx \int_{1+|x\vec{\beta}_f + (1-x)\vec{\beta}_i|}^{1-|x\vec{\beta}_f + (1-x)\vec{\beta}_i|} \frac{du}{u^2} = \frac{4\pi}{1 - |x\vec{\beta}_f + (1-x)\vec{\beta}_i|^2}$$

Note that $|x\vec{\beta}_f + (1-x)\vec{\beta}_i|^2$ contains the exit angle of the electron due to the scalar product of $\vec{\beta}_f$ and $\vec{\beta}_i$ so:

$$1 - |x\vec{\beta}_f + (1-x)\vec{\beta}_i|^2 = 1 - \beta^2 + 4x\beta^2(1-x)\sin^2\frac{\theta}{2}$$

We must now perform the integration over the variable x:

$$4\pi \int_0^1 \frac{dx}{1 - \beta^2 + 4x\beta^2(1-x)\sin^2\frac{\theta}{2}} \tag{7.32}$$

Note that the denominator is in the quadratic form $ax^2 + bx + c$ with $a = -4\beta^2\sin^2\frac{\theta}{2}$, $b = 4\beta^2\sin^2\frac{\theta}{2}$ and $c = 1 - \beta^2$. Note also that the solution for this type of integral depends on whether $4ac - b^2$ is 0, greater than 0 or less than < 0. It is in this case less than 0, and one solution (there is another in terms of ln) is:

$$\int \frac{dx}{ax^2 + bx + c} = \frac{-2}{\sqrt{b^2 - 4ac}} \operatorname{arctanh}\left(\frac{2ax + b}{\sqrt{b^2 - 4ac}}\right) + C \tag{7.33}$$

Where C is a constant of the integration.

It should be easy to verify that $\sqrt{b^2 - 4ac} = 4\beta\sin\frac{\theta}{2}\sqrt{1 - \beta^2\cos\frac{\theta}{2}}$. Using this and the integration with the limits $[1, 0]$ over x, we obtain:

$$\frac{(-2)(4\pi)}{4\beta\sin\frac{\theta}{2}\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}}\times\left[\operatorname{arctanh}\left(-\frac{\beta\sin\frac{\theta}{2}}{\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}}\right)-\operatorname{arctanh}\left(\frac{\beta\sin\frac{\theta}{2}}{\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}}\right)\right] \tag{7.34}$$

The factor of (4π) has been brought forward from equation Eqn. 7.32. We will examine 7.34 in two limits: the non-relativistic ($\beta << 1$) and ultra-relativistic ($\beta \approx 1$). We begin by first utilizing the Taylor expansion for arctanh(x) for $|x| < 1$:

$$\operatorname{arctanh}(x) = \sum_{n=0}^{\infty}\frac{x^{2n+1}}{2n+1} \tag{7.35}$$

Expanding (7.34) using (7.35) we obtain:

$$\frac{-4\pi}{\beta\sin\frac{\theta}{2}}\left[\frac{-\beta\sin\frac{\theta}{2}}{(1-\beta^2\cos^2\frac{\theta}{2})}-\frac{\beta^3\sin^3\frac{\theta}{2}}{3(1-\beta^2\cos^2\frac{\theta}{2})^2}+\cdots\right] \tag{7.36}$$

We now expand out the $(1-\beta^2\cos^2\frac{\theta}{2})^1$ and $(1-\beta^2\cos^2\frac{\theta}{2})^2$ terms in the denominator, dropping all terms higher than β^2 to obtain:

$$4\pi[1+\beta^2\cos^2\frac{\theta}{2}+\frac{\beta^2}{3}\sin^2\frac{\theta}{2}] = 4\pi[1+\beta^2-\frac{2\beta^2}{3}\sin^2\frac{\theta}{2}] \tag{7.37}$$

Using the final result of Eqn. 7.37 in Eqn. 7.31, and recalling the two contributions of -4π from each of the m^2 terms in Eqn 7.31, and not going above $O(\beta^2)$ and using $\cos\theta = 1-2\sin^2\frac{\theta}{2}$ in Eqn. 7.31 we obtain:

$$\frac{d\sigma}{d\Omega_f} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{ELASTIC}}\frac{e^2}{\pi}\ln\left(\frac{w_{\text{MIN}}}{w_{\text{MAX}}}\right)\times\frac{8\pi}{3}\sin^2\frac{\theta}{2} \tag{7.38}$$

This is differential cross section of Bremmstrahlung in the soft photon limit.

To calculate the cross-section in the relativistic limit, i.e $\beta \approx 1$, we use the following identity in order to obtain the desired form:

$$\operatorname{arctanh} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \tag{7.39}$$

We substitute approximately for $\operatorname{arctanh} y$ in Eqn (7.33) and obtain after some straightforward simplification:

$$\frac{2\pi}{\beta\sin\frac{\theta}{2}\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}}\left[\ln\left(\frac{\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}+\beta\sin\frac{\theta}{2}}{\sqrt{1-\beta^2\cos^2\frac{\theta}{2}}-\beta\sin\frac{\theta}{2}}\right)\right] \tag{7.40}$$

Now as $\beta\to 1$, we have to be careful about the *denominator* of the argument of the log. In all other places we will simply take the limit $\beta = 1$.

In the denominator of the argument of the log in Eqn. 7.40 we let $\beta = 1-\delta$

$$\sqrt{1-(1-\delta)^2\cos^2\frac{\theta}{2}}-(1-\delta)\sin\frac{\theta}{2}$$

ignoring terms of $O(\delta^2)$ and higher we obtain:

$$\sin\frac{\theta}{2}\sqrt{1+2\delta\cot^2\frac{\theta}{2}}-(1-\delta)\sin\frac{\theta}{2}$$

$$\sin\frac{\theta}{2}(1+\delta\cot^2\frac{\theta}{2})-\sin\frac{\theta}{2}+\delta\sin\frac{\theta}{2}=\frac{\delta}{\sin\frac{\theta}{2}}$$

Inserting this back, approximately into Eqn. 7.40 after setting $\beta = 1$ in all other terms we get

$$\frac{2\pi}{\sin^2\frac{\theta}{2}}\cdot\ln\left(\frac{2\sin^2\frac{\theta}{2}}{\delta}\right) \tag{7.41}$$

We now have to think about obtaining an expression for δ in terms of known, measurable quantities. To do this, we first assume that this process is very nearly elastic (soft photon approximation) in addition to being ultra-relativistic. This means that $E_i \approx E_f$, $|\vec{p}_i| \approx |\vec{p}_f|$ and that $|\vec{p}_i| \approx |\vec{p}_f| \approx E_i \approx E_f \approx E$.

We note that with these assumptions, the four momentum transfer can be simplified. We begin with:

$$q^2 = (p_f - p_i)^2 = 2m^2 - 2(E^2 - E^2\cos\theta)$$

Note now that in the ultra-relativistic limit, $E >> m$. Using this fact and $\cos\theta = 1-2\sin^2\frac{\theta}{2}$, we get:

$$-q^2 = 4E^2\sin^2\frac{\theta}{2}$$

using $E^2 = m^2\gamma^2 = \frac{m^2}{(1-\beta^2)} = \frac{m^2}{(1-(1-\delta)^2)} \simeq \frac{m^2}{2\delta}$ we obtain

$$\frac{-q^2}{m^2} = \frac{2\sin^2\frac{\theta}{2}}{\delta} \tag{7.42}$$

Using Eqn. 7.42 in Eqn. 7.41 we obtain:

$$\frac{d\Omega_k}{(1-\hat{k}\cdot\vec{\beta}_i)(1-\hat{k}\cdot\vec{\beta}_f)} = \frac{2\pi}{\sin^2\frac{\theta}{2}} \times \ln\left(\frac{-q^2}{m^2}\right) \tag{7.43}$$

This in turn can be used in Eqn. 7.31. We also recall that

$$2(1-\beta^2\cos\theta) \approx 4\sin^2\frac{\theta}{2}$$

and that the two terms containing $-\frac{m^2}{E_{f,i}}$ contribute -4π each. Doing all of this we obtain:

$$\frac{d\sigma}{d\Omega_f} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{ELASTIC}} \times \frac{2e^2}{\pi} \times \ln\left(\frac{\omega_{\text{MIN}}}{\omega_{\text{MAX}}}\right) \times \left[\ln\left(\frac{-q^2}{m^2}\right) - 1\right] \tag{7.44}$$

This is the differential cross section for Bremmstrahlung in the ultra-relativistic, soft photon limit.

(1) Physics comments on Bremsstrahlung
(2) Screening by Coulomb field of electrons
(3) Speak to Chris about quick interpretation.
(4) Problem: Simply redo for two photon emission.
(5) Feynmann graph development ask reader to redo and verify that M is same

REGARDING FEYNMANN GRAPH DEVELOPMENT : This is perhaps a slightly odd place to put this information but is as good as any. The reader should be reminded in oppurnistically:

1 Note that a delta function in energy appears when a scattering is off an immovable object (Mott scattering). This invokes a factor T in the square of the amplitude. This also means a factor of only 1/V arising from the square of the incoming particles wave, the current always has a 1/V factor in it no matter how many particle come in. Each counting factor has a factor of V in the numerator.

2 For the potential from a moving particle we get a 4-d delta function the square of whihc provides the T for the rate and an extra factor of V. For two incoming particles one of theese Vs cancels the V of one one the

incoming particle wave function box-norms, the remaining incoming particle box norm V is canceld by the 1/V from the J. The outgoing particle box norm Vs are canceled exactly by the counting factor Vs of which we must by definition have the same number. (Norms are squared)

3 Potentials and delta functions. Internal lines.

7.1 Pair Creation of e^+e^- by a photon in a Coulomb Field

In the chapter on propagators and scattering the narrative of hole theory and the properties of our propagators were used to describe the *experimentally observed* creation of electron and positron pairs. (REFERENCE AND MODIFICATION OF PREVIOUS TEXT NEEDED). Given that negative energy states are always filled we imagined a sufficiently strong potential (read a sufficiently energetic incoming photon) knocking out an electron from this sea and the who and electron are the electron positron pair. We frame the steps in the scattering leading to the emergence of an e^+e^- pair:

(1) We begin with a *negative energy* electron at $t' \to \infty$. Due to our propagator it will move backward in time thus appearing as an e^+ moving forward in time at $t' = +\infty$
(2) As it moves backward in time this state is scattered by a potential (remember the potential can mix $E < 0$ and $E > 0$ states) into an $E > 0$ state which is propagated forward in time and observed at $t = \infty$ as an e^-
(3) Comparing (1) and (2) [NOTE FROM REBECCA: (1) AND (2) OF WHICH SECTION?] we note the emergence of an e^+e^- pair at $t = \infty$.

(CHECK NARRATIVE: LEFTOVER: SOUNDS DODGY).

We now write down the perturbative series for the S-matrix:

$$S_{fi} = -\delta_{fi} - ie \int d^4x_1 \bar{\psi}_f(x_1) \not{A}(x_1)\psi_i(x_1) - ie^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_f(x_2) \not{A}(x_2) S_F(x_2 - x_1) \not{A}(x_1)\psi_i(x_1) \tag{7.45}$$

[LEFTOVER: cogent explanation of $-ie^n$ instead of $+ie$. Greiner page 56-57 is good. $t' \to \infty$, $\theta(t' - t)$]

Per our narrative $\psi_i(x)$ is a negative energy state as $t' \to \infty$ and $\bar{\psi}_f(x)$ is a positive energy state. We will not explicitly insert the potential as done for previous amplitudes as we shall see it suffices to use Bremstrahlung and crossing symmetry.

Consider now an incoming photon with 4-momentum $k = (w, \vec{k})$ and the 4-momenta of the electron and positron $p_- = (E_-, \vec{p}_-)$ and $p_+ = (E_+, \vec{p}_+)$ respectively. Then we can write:

$$k^2 = \omega^2 - |\vec{k}|^2 = 0 = 2m^2 + 2E_+E_- - 2\vec{p}_+ \cdot \vec{p}_-$$

If we view the process in the center of mass frame $|\vec{p}_+| = |\vec{p}_-|$ etc:

$$0 = 2m^2 + 2|\vec{p}|^2 + wm^2 + 2|\vec{p}|^2$$

$$0 = 4m^2 + 4|\vec{p}|^2$$

this condition can obviously never be satisfied, another agent is required to absorb the appropriate energy momentum. This agent is the Coulombic potential of the nucleus. We observe now that the same arguments as we invoked in Bremmstrahlung apply. This will be a second order process, using crossing symmetry we make the following changes to the Bremmstrahlung amplitue:

(1) $\bar{u}_f \to \bar{u}_-$ (symbolic change for outgoing elcetron)
 $u_i \to v_+$ (incoming electron changes to outgoing or incoming positron (?))
(2) $p_i \to -p_+$ (positron momentum reversed)
 $p_f \to p_-$ (change of symbol for final electron momentum)
(3) $k \to -k$ (photon momentum reversed: outgoing becomes incoming).

After these substitutions one obtains for just the matrix element

$$M_{fi} = \frac{iZe^3 4\pi}{|\vec{q}|^2} \bar{u}_- \left[\frac{\gamma^0(-\not{p}_+ + \not{k} + m)\not{\epsilon}}{(-p_+ + k)^2 - m^2 + i\epsilon} + \frac{\not{\epsilon}(\not{p}_- - \not{k} + m)\gamma^0}{(E-k)^2 - m^2 + i\epsilon} \right] v_+ \tag{7.46}$$

(LEFTOVER: check overall minus sign).

Squaring the matrix element, summing over electron and positron spins and averaging over the (incoming) photon polarizations (ϵ) we obtain with indices:

$$|M_{fi}|^2 = \frac{Z^2e^6(4\pi)^2}{|\vec{q}|^4 2} \left(\frac{\not{p}_- + m}{2m}\right)_{\alpha\rho} [\mathbb{M}]_{\alpha\beta} \left(\frac{\not{p}_+ - m}{2m}\right)_{\beta\sigma} [\mathbb{M}^*]_{\sigma\rho} \tag{7.47}$$

where SPIN denotes the sum over e^+e^- spins *and* the average over photon polarizations.

Where the following substitutions have been made:

$$\mathbb{M} = \frac{g^0(-\not{p}_+ + \not{k} + m)\epsilon}{(-\not{p}_+ + k)^2 - m^2 + i\epsilon} + \frac{\not{\epsilon}(\not{p}_- - \not{k} + m)\gamma^0}{(p_- - k)^2 - m^2 + i\epsilon}$$

$\mathbb{M}^*$ is the complex conjugate of $\mathbb{M}$

$$\sum_{S_-} u_-\bar{u}_- = \frac{(\not{p}_- - m)}{2m} \text{ and } \sum_{S_+} v_+\bar{v}_+ = \frac{(\not{p}_+ - m)}{2m}$$

with $\vec{q} = \vec{p}_+ + \vec{p}_- - \vec{k}$

$$|M_{fi}|^2_{SPIN} = \frac{Z^2 e^6 (4\pi)^2}{|\vec{q}|^4 8m^2} Tr[(\not{p}_- + m)\mathbb{M}(\not{p}_+ - m)\mathbb{M}] \tag{7.48}$$

[Note to self: think about using $\sum_{S_+} v\bar{v}_+ = -\frac{(-p_+ + m)}{2m}$, seems to be the norm and for some reason makes it easier to keep track of the minus sign.]

It's easy to see that the simplification of the trace term is going to include several terms of traces of up to 8 γ matrices. The manipulation of these is not very educational. Note also that the limit of soft photon, k 0 cannot be taken here since the photon must be energetic enough to produce an electron positron pair thus $|\vec{k}| = w > 2m$.

The full amplitude modulus squared with all the plane wave and counting factors of the outgoing states and converted δ function in energy inserted (REFERENCE TO BREMSS + PERHAPS MOTT) gives us: $(|M_{fi}|^2_{SPIN} = \frac{1}{2}\sum_{\epsilon_\gamma}\sum_{S_+S_-}|M_{fi}|^2)$

$$|S_{fi}|^2 = \frac{2\pi}{wV}\frac{m}{E_-V}\frac{m}{E_+V}(T2\pi\delta(E_+ + E_- - \omega))|M_{fi}|^2_{SPIN}\frac{Vd^3\vec{p}_+}{(2\pi)^3}\frac{Vd^3\vec{p}_-}{(2\pi)^3} \tag{7.49}$$

the factors of $\frac{2\pi}{wV}$ $\frac{m}{E_\pm V}$ are the squares of the box-normalization factors of the photon (incoming) and electron and positron (outgoing) wave functions. The factor of $T2\pi\delta(E_+ + E_- - w)$ is the square of the Dirac δ function in *energy* first encountered in Chapter 5 (reference needed). Dividing by T on both sides to convert to a rate and by the incident flux of the incoming photon we obtain the following expression for the differential cross section:

$$d\sigma = \frac{dR}{J_\gamma} = \frac{|S_{fi}|^2}{TJ_\gamma} = \frac{(2\pi)^2 m^2}{wE_-E_+}\delta(E_+ + E_- - \omega)|M_{fi}|^2_{SPIN}\frac{d^3\vec{p}_+}{(2\pi)^3}\frac{d^3\vec{p}_-}{(2\pi)^3} \tag{7.50}$$

where the cross-section is left incomplete. a more thorough treatment is given here: (REFERENCE NEEDED), $J_{\gamma,INC}$ is simply: $\frac{|\vec{V}|}{V} = \frac{c}{V} = \frac{1}{V}$ (LEFTOVER CHECK)

Simplifying the previous expression we obtain:

$$d\sigma = \frac{Z^2 e^6}{8\pi\omega E_+ E_-}\frac{1}{|\vec{q}|^4}\times\delta(E_+ + E_- - \omega)Tr[(\not{p}_- + m)\mathbb{M}(\not{p}_+ - m)\times\mathbb{M}^*]d^3\vec{p}_- d^3\vec{p}_+ \tag{7.51}$$

here $\mathbb{M}$ has been defined in Eqn. 7.47. We note that $\mathbb{M}$ has two terms that contain the factors from the momentum space electron and positron propagators these are $\frac{1}{(-p_+ + k)^2 - m^2 + i\epsilon}$ and $\frac{1}{(p_- - k)^2 - m^2 + i\epsilon}$

The denominators will simplify to $\frac{1}{-2p_+\cdot k}$ and $\frac{1}{-2p_-\cdot k}$ respectively. The scalar product with k of the positron and electron momenta contain an *overall* factor of $|\vec{k}| = w$ (for example $p_+\cdot k = \omega E_+ - \vec{k}\cdot\vec{p}_+ = \omega(E_+ - \hat{k}\cdot\vec{p}_+)$, the presence of $\mathbb{M}$ and $\mathbb{M}^*$ in the trace means that we can factorize a factor of $\frac{1}{\omega^2}$ from the trace term, appropriately replacing $\mathbb{M}$, $\mathbb{M}^*$ by $\mathbb{M}'$, $\mathbb{M}'^*$ where the redefinition should be obvious we can re-write the differential cross-section in Eqn. 7.51 to show its behaviour with respect to the incoming poton energy:

$$d\sigma = \frac{Z^2 e^6}{8\pi^2\omega^3 E_+ E_-}\times\frac{1}{|\vec{q}|^4}\delta(E_+ + E_- - \omega)Tr[(\not{p}_- + m)\mathbb{M}'(\not{p}_+ - m)\mathbb{M}'^*]\times d^3\vec{p}_+ d^3\vec{p}_- \tag{7.52}$$

Using $d^3\vec{p}_\pm = |\vec{p}_\pm|E_\pm dE_\pm d\Omega_\pm$ we can integrate over the electron energy to obtain:

$$d\sigma = \frac{Z^2 e^6}{8\pi^2\omega^2}\frac{|\vec{p}_+||\vec{p}_-|}{|\vec{q}|^4}Tr[(\not{p}_- + m)\mathbb{M}'(\not{p}_+ - m)\mathbb{M}'^*]\times dE_+ d\Omega_+ d\Omega_- \tag{7.53}$$

Eqn (7.53) must be multiplied by $\theta(w - E_+ - m)$, since we have integrated over all electron energies and want Eqn. 7.53 to represent the production of both an e^+ and e^- we account for the electron to be stationary at creation. As mentioned earlier the traces in Eqn. 7.53 are horrendous to evaluate and so we leave the differential cross-section as it is. We also note that terms in $\mathbb{M}'$ and $\mathbb{M}'^*$ contain $|\vec{p}_-|$ in the denominator which can cancel out the $|\vec{p}_-|$ in the numerator (LEFTOVER: REDESCRIBE) for the limit $|\vec{p}_-| \cong 0$. This a lowest possible order calculation, in reality the interaction of the e^+e^- with the Coulombic potential of the nucleus must be also taken into account, the validity of Eqn. 7.53 is thus limited. (REFERENCE NEEDED: GREINER)

Chapter 8

Compton Scattering

Compton scattering involves the scattering of free photons in matter, the change in the energy of the photon, semiclassically, gives us a scattered photon with a changed wavelength. We will now consider this process using our relativistic quantum mechanical framework, interpreting this as a process in which the change in momentum of the struck electron (acceleration) due to the incident photon, causes it to radiate another photon. It should be easy for the reader to verify that this process can only take place at second order, analogous to the argument presented at the beginning of the section on Bremsstrahlung. We are a little bit wiser now and a recollection of the Bremsstrahlung calculation will lead us to conclude that we need to evaluate the sum of the potentials of the incident and scattered photons at x_1 and x_2 in the S_{fi} expansion, and to denote the plane wave associated with the outgoing photon by $\epsilon^{\mu\prime}\sqrt{\frac{2\pi}{\omega' V}}e^{ik'\cdot x}$ and the incoming by $\epsilon^{\mu}\sqrt{\frac{2\pi}{\omega V}}e^{-ik\cdot x}$. Using these and taking care to isolate only those terms with *one of each* of these potentials, we have:

$$
-ie^2\sqrt{\frac{m^2}{E_f E_i}}\frac{2\pi}{\sqrt{\omega\omega'}}\frac{1}{V^2}\int\frac{d^4p}{(2\pi)^4}\int d^4x_1\int d^4x_2\bar{u}_f e^{ip_f\cdot x_2}
$$

$$
\times\left[\frac{\not{\epsilon}'(\not{p}+m)\not{\epsilon}e^{ik'\cdot x_2-ik\cdot x_1}}{p^2-m^2+i\epsilon}+\frac{\not{\epsilon}(\not{p}+m)\not{\epsilon}'e^{ik'\cdot x_1-ik\cdot x_2}}{p^2-m^2+i\epsilon}\right]
$$

$$
\times e^{-ip\cdot(x_2-x_1)}\cdot e^{-ip_i\cdot x_1}u_f
$$

where (ω, k), (ω', k') represent the 4-momenta of the incoming and outgoing momenta. Note that both orderings in time of the incident and radiated photons will be contribute. The reader should verify that the integrals over the dummy variables yield δ functions and one of the δ functions can be integrated over $\frac{d^4p}{(2\pi)^4}$ to fix the momentum of the propagator term to $p_i + k$ and $p_i - k'$ in the two amplitudes. It is straightforward to verify that the transition amplitude is:

$$S_{fi} = \sqrt{\frac{m^2}{E_f E_i \omega\omega'}} \cdot \frac{2\pi}{V^2}(2\pi)^4 \delta(p_f - p_i + k' - k) M_{fi} \tag{8.1}$$

where M_{fi} is given by:

$$M_{fi} = -ie^2 \bar{u}_f \left[\frac{\not{\epsilon}'(\not{p}_i + k + m)\not{\epsilon}}{(p_i + k)^2 - m^2 + i\epsilon} + \frac{\not{\epsilon}(\not{p}_i - \not{k}' + m)\not{\epsilon}'}{(p_i - k')^2 - m^2 + i\epsilon} \right] u_i \tag{8.2}$$

Recall from the Bremsstrahlung calculation that $(p_i - k)^2 = 2p_i \cdot k$, $(p_i - k')^2 = -2p_i \cdot k'$. We can also utilize the Dirac equation $(\not{p}_i - m)u_i = 0$, by putting $\not{\epsilon}$ and $\not{\epsilon}'$ to the left of $\not{p}_i$ in the two amplitudes in Eqn. 8.3. Doing all of will give us terms of $2\not{\epsilon}' p_i \cdot \epsilon$ and $2\not{\epsilon} p_i \cdot \epsilon'$ in the numerator. It's easy to carry out these manipulations and obtain:

$$S_{fi} = -ie^2 \sqrt{\frac{m^2}{E_f E_i \omega\omega'}} \frac{2\pi}{V^2}(2\pi)^4 \delta^4(p_f - p_i + k - k) \tag{8.3}$$

$$\times \bar{u}_f \left[\frac{(2\not{\epsilon}' p_i \cdot \epsilon + \not{\epsilon}' \not{k} \not{\epsilon})}{2p_i \cdot k} - \frac{(2\not{\epsilon} p_i \cdot \epsilon' - \not{\epsilon} \not{k}' \not{\epsilon}')}{2p_i \cdot k'} \right] u_i$$

As discussed earlier, $\partial_\mu A^\mu = 0$ and with $A_\mu = \epsilon_\mu e^{ik \cdot x}$, this gives us $\epsilon \cdot k = \epsilon' \cdot k' = 0$, since $k \cdot k = k' \cdot k' = 0$. This in turn means that we can modify ϵ and ϵ' to $\epsilon + \alpha k$ and $\epsilon' + \beta k'$ where α and β are arbitrary constants. We will utilize this gauge freedom in the following steps.

We can use this freedom to redefine our polarization vectors to the following (as we shall see) convenient choice:

$$\epsilon_\mu \to \epsilon_\mu - \frac{(p_i \cdot \epsilon)}{(p_i \cdot k)} k_\mu$$

$$\epsilon'_\mu \to \epsilon'_\mu - \frac{(p_i \cdot \epsilon')}{(p_i \cdot k')} k'_\mu$$

(Reference needed Greiner or standard)

With this choice, $p_i \cdot \epsilon = p_i \cdot \epsilon' = 0$ and as before $\epsilon \cdot k = \epsilon' \cdot k' = 0$, since $k \cdot k = k' \cdot k' = 0$. Our amplitude now becomes:

$$S_{fi} = -ie^2\sqrt{\frac{m^2}{E_f E_i \omega\omega'}} \times \frac{2\pi}{V^2}(2\pi)^4\delta^4(p_f - p_i + k' - k)\bar{u}_f \left[\frac{\not{\epsilon}'\not{k}\not{\epsilon}}{2p_i \cdot k} + \frac{\not{\epsilon}\not{k}'\not{\epsilon}'}{2p_i \cdot k'}\right] u_i \tag{8.4}$$

We will now have to modulus square this and insert the usual flux and state density counting factors. The most complicated piece of our remaining procedure is of course the squaring of

$$\bar{u}_f \left[\frac{\not{\epsilon}'\not{k}\not{\epsilon}}{2p_i \cdot k} + \frac{\not{\epsilon}\not{k}'\not{\epsilon}'}{2p_i \cdot k'}\right] u_i$$

and summing and averaging over the spins as always, Thus we need to sum and average:

$$\left(\bar{u}_f \left[\frac{\not{\epsilon}'\not{k}\not{\epsilon}}{2p_i \cdot k} + \frac{\not{\epsilon}\not{k}\not{\epsilon}'}{2p_i \cdot k'}\right] u_i\right) \left(\bar{u}_f \left[\frac{\not{\epsilon}'\not{k}\not{\epsilon}}{2p_i \cdot k} + \frac{\not{\epsilon}\not{k}'\not{\epsilon}'}{2p_i \cdot k'}\right] u_i\right)^* \tag{8.5}$$

We first need to understand the complex conjugate of 3 gamma matrices between the $\bar{u}_f$ and u_i spinors. Writing this out explicitly with all indices present:

$$(\bar{u}_{f\sigma}\gamma^{\alpha}_{\sigma\rho}\gamma^{\beta}_{\rho\lambda}\gamma^{\delta}_{\lambda\phi}u_{i\phi})^* = u^{+}_{f\chi}\gamma^{0}_{\chi\sigma}\gamma^{\alpha}_{\sigma\rho}\gamma^{\beta}_{\rho\lambda}\gamma^{\delta}_{\lambda\phi}u_{i\phi})^* \tag{8.6}$$

We have to take the complex conjugate of each piece in the brackets.

Note that

(1) The complex conjugate of $u_{i\phi}$ is simply $u^{+}_{i\phi}$ or the ϕ^{th} element of u^{+}_{i}
(2) The complex conjugate of $u^{+}_{f\chi}$ is $u_{f\chi}$ or the χ^{th} element of u_f.
(3) γ^0 is unaffected
(4) The complex conjugates of the elements of the γ matrices are simply the elements of the transposes of their Hermitian conjugates. Thus for example $\gamma^{\alpha *}_{\sigma\rho} = \gamma^{\alpha +}_{\rho\sigma}$ and of course $\gamma^{\mu +} = \gamma^0\gamma^\mu\gamma^0$. Thus

$$\gamma^{\alpha *}_{\sigma\rho} = \gamma^{0}_{\rho z}\gamma\Sigma\eta^{\alpha}\gamma^{0}_{\eta\sigma} \tag{8.7}$$

Recognizing that

$$(u^{+}_{f\chi}\gamma^{0}_{\lambda\sigma}\gamma^{\alpha}_{\sigma\rho}\gamma^{\beta}_{\rho\lambda}\gamma^{\delta}_{\lambda\phi}u_{i\phi})^* = (u_{f\chi}\gamma^{0}_{\lambda\sigma}\gamma^{\alpha +}_{\sigma\rho}\gamma^{\beta +}_{\rho\lambda}\gamma^{\delta +}_{\lambda\phi}u^{+}_{i\phi}) \tag{8.8}$$

and using relations like (8.5) to "sandwich" each of the γ^μs in (8.8) it is easy to show that (8.8) equals $u^+_{i\phi}\gamma^0_{\phi\Sigma}\gamma^\delta_{\Sigma\theta}\gamma^\beta_{\theta\rho}\gamma^\alpha_{\rho\chi}u_f$.

This simply means that the second term (on the right in the brackets) in Eqn. 8.5 can be rewritten and Eqn. 8.5 becomes

$$\bar{u}_f\left[\frac{\not\epsilon'\not k\not\epsilon}{2p_i\cdot k}+\frac{\not\epsilon\not k'\not\epsilon'}{2p_i\cdot k'}\right]u_i\bar{u}_i\left[\frac{\not\epsilon'\not k'\not\epsilon}{2p_i\cdot k'}+\frac{\not\epsilon\not k\not\epsilon'}{2p_i\cdot k}\right]u_f \tag{8.9}$$

This must be multiplied by $\frac{1}{2}$ and summed over all spins.

The reader who has diligently followed all the scattering calculations so far can verify that $\frac{1}{2}\sum_{S_i,S_f}$ overEqn. 8.9, will yield:

$$\frac{1}{32m^2}Tr\left[(\not p_f+m)\left(\frac{\not\epsilon\not k\not\epsilon}{p_i\cdot k}+\frac{\not\epsilon\not k'\not\epsilon}{p_i\cdot k'}\right)(\not p_i+m)\left(\frac{\not\epsilon'\not k'\epsilon}{p_i\cdot k'}+\frac{\not\epsilon\not k\not\epsilon'}{p_i\cdot k}\right)\right]$$

Multiplying everything out in the square brackets we collect terms with denominators $(p_i\cdot k)^2$, $(p_i\cdot k')^2$ and $(p_i\cdot k)(p_i\cdot k')$, and deal with each separately.

The reader is reminded that the following will serve to make our calculation easier:

(1) $\epsilon\cdot\epsilon = 1$, $\epsilon'\cdot\epsilon' = -1$, $k\cdot k = 0$, $k'\cdot k' = 0$, $\epsilon\cdot k = 0$, $\epsilon'\cdot k' = 0$, $\epsilon\cdot p_i = \epsilon'\cdot p_i = 0$

(2) All the trace theorems in general and in particular

$$Tr(\not a\not b\not c\not d\dots) = 2b\cdot cTr(\not a\not d\dots) - Tr(\not a\not c\not b\not d\dots)$$

will let us move terms around.

(3) When one encounters two *like* terms in a trace ie $Tr(\not A\ \not B\ \not B\ \not C\ \not D\dots)$ then this is $B\cdot BTr(\not A\ \not C\ \not D\dots)$

(4) Finally we remind the reader of the 4-momentum conservation imposed by the Dirac δ function in Eqn. 8.4 i.e $p_f = p_i + k - k'$

Our general plan is obviously to move all ϵ and ϵ' next to each other and similarly for k and k'. Finally, some of the terms have an odd number of γ entries; needless to say these are trivially zero.

The three terms which we will treat separately are

$$\frac{1}{32m^2}Tr\left[\frac{(\not p_f+m)(\not\epsilon'\not k\not\epsilon)(\not p_i+m)(\not\epsilon\not k\not\epsilon')}{(p_i\cdot k)^2}\right] \tag{8.10}$$

$$\frac{1}{32m^2}Tr\left[\frac{(\not p_f+m)(\not\epsilon\not k'\not\epsilon')(\not p_i+m)(\not\epsilon'\not k'\not\epsilon)}{(p_i\cdot k')^2}\right] \tag{8.11}$$

and the cross term, which the reader can verify by the cyclicity of traces is simply

$$\frac{1}{32m^2} Tr\left[\frac{(\not{p}_f + m)(\not{\epsilon}\not{k}'\not{\epsilon}')(\not{p}_i + m)(\not{\epsilon}\not{k}\not{\epsilon}')}{(p_i \cdot k)(p_i \cdot k')}\right] \tag{8.12}$$

$$Tr(\not{p}_f\,\not{\epsilon}\not{k}\not{\epsilon}\not{p}_i\,\not{\epsilon}\not{k}\not{\epsilon}' + m^2\,\not{\epsilon}'\,\not{k}\not{\epsilon}\not{\epsilon}\not{k}\not{\epsilon}')$$

The second term is trivial to simplify. We observe that $\not{\epsilon}\not{\epsilon} = \epsilon \cdot \epsilon = -1$ which brings $\not{k}\not{k}$ next to each other, so $\not{k}\not{k} = k \cdot k = 0$, and we are left with only the first term (we have of course not even considered terms with odd numbers of γs)

$$Tr(\not{p}_f\,\not{\epsilon}'\,\not{k}\not{\epsilon}\not{p}_i\,\not{\epsilon}\not{k}\not{\epsilon}') = -Tr(\not{p}_f\,\not{\epsilon}'\,\not{k}\not{p}_i\,\not{\epsilon}\not{\epsilon}\not{k}\not{\epsilon}')$$

$$= Tr(\not{p}_f\,\not{\epsilon}'\,\not{k}\not{p}_i\,\not{k}\not{\epsilon}') = 2p_i \cdot k Tr(\not{P}_f\,\not{\epsilon}'\,\not{k}\not{\epsilon}')$$

$$= 2p_i \cdot k[(p_f \cdot \epsilon')Tr(\not{k}\not{\epsilon}') - p_f \cdot k Tr(\not{\epsilon}'\not{\epsilon}') + p_f \cdot \epsilon Tr(\not{\epsilon}'\,\not{k})$$

which yields after substituting $p_f = p_i + k - k'$

$$2p_i \cdot k[8(k \cdot \epsilon')^2 + 4p_i \cdot k - 4k' \cdot k] \tag{8.13}$$

Note that in the above expression (8.13), we have used relations 1-4 preceding equation (8.9). Note that $p_f = p_i - k' + k$ gives us $p_f^2 = (p_i - k' + k)^2$. Using $p_f^2 = p_i^2 + m^2$ we obtain $-2p_i \cdot k' + 2p_i \cdot k - 2k \cdot k' = 0$ which gives $p_i \cdot k' = p_i \cdot k - k' \cdot k$, which we substitute in Eqn. 8.13 to give

$$2p_i \cdot k[8(k \cdot \epsilon')^2 + 4(p_i \cdot k')] \tag{8.14}$$

where a factor of $\frac{1}{16m^2} \times \frac{1}{(p_i \cdot k)^2}$ is understood.

We now move on to expression (8.11). Expanding this we obtain

$$\frac{1}{32m^2(p_i \cdot k')^2} Tr(\not{p}_f\,\not{\epsilon}\not{k}'\,\not{\epsilon}'\,\not{p}_i\,\not{\epsilon}'\,\not{k}'\,\not{\epsilon})$$

as the only non zero term, (the reader should verify this).

We proceed in a manner similar to term (8.10) and we obtain

$$2k' \cdot p_i[4p_i \cdot k' + 4k \cdot k' - 8(k' \cdot \epsilon)^2]$$

Using $k \cdot k' = p_i \cdot k - p_i \cdot k'$ we obtain:

$$2p_i \cdot k'[4p_i \cdot k - 8(k' \cdot \epsilon)^2] \tag{8.15}$$

The reader should verify all of the steps in between.

We now deal with the mixed term (8.12).

We have:

$$\frac{1}{18m^2(p_i \cdot k)(p_i \cdot k')} Tr(\not{p}_f \not{\epsilon}' \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon}' + m^2 \not{\epsilon} \not{k}' \not{\epsilon}' \not{\epsilon} \not{k} \not{\epsilon}') \tag{8.16}$$

Droppoing the factor outside the trace for our manipulations, we use $\not{p}_f = \not{p}_i + \not{k} - \not{k}'$ to obtain

$$Tr(\not{p}_i \not{\epsilon} \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon}' + \not{k} \not{\epsilon} \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon}' - \not{k}' \not{\epsilon} \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon}') \tag{8.17}$$

To get the reader started, we consider the first term in the brackets in Eqn. 8.17 and leave the verification of the other two terms to the reader.

The first term is:

$$Tr(\not{p}_i \not{\epsilon} \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon}') = 2p_i \cdot \epsilon Tr(\not{k}' \cdots \not{\epsilon}') - Tr(\not{\epsilon} \not{p}_i \not{k}' \not{\epsilon}' \not{p}_i \not{\epsilon} \not{k} \not{\epsilon})$$

this is just zero since $p_i \cdot \epsilon = 0$.

The second term is simply:

$$-2p_i \cdot \epsilon' Tr(\not{\epsilon} \not{p}_i \not{k}' \not{\epsilon} \not{k} \not{\epsilon}') + Tr(\not{\epsilon} \not{p}_i \not{k}' \not{p}_i \not{\epsilon}' \not{\epsilon} \not{k} \not{\epsilon}')$$

and once again this a zero for the first term and the second term can be manipulated :

$$2k' \cdot p_i \cdot Tr(\not{\epsilon} \not{p}_i \not{\epsilon}' \not{\epsilon} \not{k} \not{\epsilon}') - Tr(\not{\epsilon} \not{p}_i \not{p}_i \not{\epsilon}' \not{\epsilon} \not{k}' \not{\epsilon}')$$

Since $\not{p}_i \not{p}_i = m^2$, the second term cancels the m^2 term in the trace in expression 8.16. The term $2k' \cdot p_i Tr(\not{\epsilon} \not{p}_i \not{\epsilon}' \not{\epsilon} \not{k} \not{\epsilon})$, after similar manipulations, will yield

$$16(p_i \cdot k')(p_i \cdot k)(\epsilon \cdot \epsilon')^2 - 8(p_i \cdot k')(p_i \cdot k) \tag{8.18}$$

The terms $\not k \not\epsilon \not k' \not\epsilon' \not p_i \not\epsilon \not k \not\epsilon'$ and $-\not k' \not\epsilon \not k' \not\epsilon' \not p_i \not\epsilon \not k \not\epsilon'$ yield $-8(\epsilon' \cdot k)^2 \cdot p_i \cdot k'$ and $+8(\epsilon \cdot k') \cdot (p_i \cdot k)$ respectively and the reader is encouraged to verify this. (18) and (19)

The cross term is then:

$$16(p_i \cdot k')(p_i \cdot k)(\epsilon \cdot \epsilon')^2 - 8(p_i \cdot k')(p_i \cdot k) - 8(\epsilon' \cdot k)^2 (p_i \cdot k) + 8(\epsilon \cdot k')(p_i \cdot k) \quad (8.19)$$

Collecting (8.19), (8.14) and (8.15) and reinserting the denominators, we have for our trace terms:

$$\frac{1}{32m^2}\left[\frac{2(p_i \cdot k)[8(k \cdot \epsilon')^2 + 4(p_i \cdot k')]}{(p_i \cdot k)^2} + \frac{2(k' \cdot p_i)[4(p_i \cdot k) - 8(k' \cdot \epsilon)^2]}{(p_i \cdot k')^2}\right] \quad (8.20)$$

$$+2\left\{\frac{16(p_i \cdot k')(p_i \cdot k)(\epsilon \cdot \epsilon')^2}{(p_i \cdot k)(p_i \cdot k')}\right.$$

$$\left. - \frac{8(p_i \cdot k')(p_i \cdot k) - 8(\epsilon' \cdot k)^2 (p_i \cdot k') + 8(\epsilon \cdot k')^2 (p_i \cdot k)}{(p_i \cdot k)(p_i \cdot k')}\right\}$$

It is a straightforward task to simplify this expression, we obtain:

$$\frac{1}{2}\sum_{S_f, s_i} |M_{fi}|^2 = \frac{e^4}{32m^4}\left[\frac{8(p_i \cdot k')}{(p_i \cdot k)} + \frac{8(p_i \cdot k)}{(p_i \cdot k')} + 32(\epsilon \cdot \epsilon')^2 - 16\right] \quad (8.21)$$

In this expression the requirement of four momentum conservation imposed by the δ function has been taken into account. One more modification is at hand: we assume the initial electron is at rest, thus $p_i = (m, 0)$ with $k = (\omega, \vec{k})$ and $k' = (\omega', \vec{k}')$. We obtain the following expression for the trace (spin sum) (denoted by $\frac{1}{2}\sum_{S_f, S_i} |M_{fi}|^2$

$$\frac{1}{2}\sum_{S_f, S_i} |M_{fi}|^2 = \frac{e^4}{32m^4}\left[\frac{8w'}{w} + \frac{8w}{w'} + 32(\epsilon \cdot \epsilon')^2 - 16\right] \quad (8.22)$$

We now proceed to complete the cross section calculation with the appropriate insertion of flux and counting factors. Recall equation (8.4), we have just completed the sum of spins of the square of the term sandwiched between and including the spinors $\bar{u}_f$ and u_i. The entire expression for $|S_{fi}|^2$, including the factor we left out, is:

$$|S_{fi}|^2 = \frac{e^4 m^2}{E_f E_i \omega' \omega} \frac{(2\pi)^2}{V^4} \left[(2\pi)^4 \delta^4(p_f - p_i + k' - k)\right]^2 \frac{1}{2} \sum_{S_f, S_i} |M_{fi}|^2 \quad (8.23)$$

From the discussion preceding and including Eqn. 5.22 we know the term in the square brackets is simply $(2\pi)^4 VT\delta(p_f - p_i + k' - k)$. Following the usual procedure, we divide by T to obtain a rate and then:

(1) Insert the counting factors for the possible outgoing electron and photon states: $\frac{Vd^3p_f}{(2\pi)^3}$, $\frac{Vd^3k'}{(2\pi)^3}$ (this should be obvious to the reader by now)
(2) Divide by the flux factor

$$\vec{J}_{INCIDENT} = \frac{|\vec{v}_1 - \vec{v}_2|}{V} = \frac{1}{E_1 E_2} \times \frac{\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}}{V}$$

We will make some general statements about the expression we obtain so for this purpose we will set $E_1 = E_i$, $E_2 = w$ but not simplify the square root immediately:

$$d\sigma = \frac{e^4 m^2 (2\pi)^2}{E_f \omega' \sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} \times (2\pi)^4 \delta^4(p_f - p_i + k' - k) \quad (8.24)$$
$$\times \frac{1}{2} \sum_{S_f, S_i} |M_{fi}|^2 \frac{d^3 p_f}{(2\pi)^3} d^3 k' (2\pi)^3$$

Recall the technique outlined in Chapter **??** to integrate out the 4-dimensional δ function over the three dimensional final momentum state differentials,

$$\frac{d^3 p}{2E} = \theta(p_0) \delta(p^2 - m^2) d^4 p$$

which introduces a 4-dimensional integral naturally, $(p_0 = E)$. In our expression (8.24) we have factors of E_f and w' in the denominator. The factor of 2 is missing so we introduce it in the denominator and the numerator. We obtain:

$$d\sigma = \frac{e^2 m^2 (2\pi)^2 \cdot 4}{\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} (2\pi)^4 \delta^4(p_f - p_i + k' - k) \quad (8.25)$$
$$\times \frac{1}{2} \sum_{S_f, S_i} |M_{fi}|^2 \frac{d^3 p}{2E_f (2\pi)^3} \frac{d^3 k'}{2w' (2\pi)^3}$$

Note that all volume factors have cancelled. Note also that the factor of m^2 comes from one incoming and one outgoing electron, the $(2\pi)^2$ from one incoming and one outgoing photon. We can now begin to build a set of rules based on experience of several scattering processes, since we've seen both electrons and photons incident and scattering we leave the reader to verify that all the factors related to flux, and counting and momentum conservation can be computed from the following set of rules. INSERT BATCH COMPTON SCATTERING PAGE 11. The reader is encouraged to verify that the expression (8.25) or any other expression in scattering processes can be built up from these.

Do appropriate check for $d\sigma$. (Note to self).

Back to finishing up, we now use $p_1 = p_i = (m, 0)$, $p_2 = (\omega, \vec{k})$ and $m_2 = 0$ in expression (8.25) to obtain:

$$d\sigma = \frac{4e^4 m}{w} \delta^4(p_f - p_i + k' - k) \left[\frac{1}{2} \sum_{S_f, S_i} |M_{fi}|^2 \right] (p_f^2 - m^2) d^4 p_f \frac{d^3 k'}{2w'} \quad (8.26)$$

Integrating over the 4-dimensional δ function we obtain:

$$d\sigma = \frac{4e^4 m}{w} \left[\frac{1}{2} \sum_{S_f, S_i} |M_{fi}|^2 \right] \theta(p_i + k - k') \delta(2mw - 2mw' - 2\vec{k} \cdot \vec{k}') \times \frac{d^3 \vec{k}'}{2w'} \quad (8.27)$$

Assuming that the θ function is satisfied, we rewrite

$$\frac{d^3 \vec{k}'}{2w'} = \frac{w'^2 dw' d\Omega'}{2w'}$$

since $|k'| = w'$ and the solid angle refers to the exit of the photon.

We next integrate over all final photon energies, incurring a factor of

$$\frac{1}{|f'(\omega')|} = \frac{1}{2m \left(1 + \frac{w}{m} \cos\theta\right)}$$

where $f(\omega')$ is the argument of the δ function. We note that the zero of f' occurs at

$$w' = \frac{w}{\left(1 + \frac{w}{m} \cos\theta\right)} \quad (8.28)$$

This result is obtainable from old quantum theory as well. We now have:

$$\frac{d\sigma}{d\Omega'} = \frac{4e^4}{2w^2}\frac{w'}{\left(1+\frac{w}{m}\cos\theta\right)} \times \frac{1}{2}\sum_{S_f,S_i}|M_{fi}|^2 \tag{8.29}$$

Using Eqn 8.28:

$$\frac{d\sigma}{d\Omega'} = \frac{2e^4 w'^2}{2w^2} \times \frac{1}{2}\sum_{S_f,S_i}|M_{fi}|^2$$

Inserting (8.27) for $\frac{1}{2}\sum_{S_f,S_i}|M_{fi}|^2$ we have:

$$\frac{d\sigma}{d\Omega'} = e^4\frac{w'^2}{w^2} \times \frac{1}{32m^2}\left[\frac{8w'}{w} + \frac{8w}{w'} + ew(\epsilon'\cdot\epsilon)^2 - 16\right]$$

$$\frac{d\sigma}{d\Omega'} = e^4\frac{w'^2}{w^2}\frac{1}{4m^2}\left[\frac{w'}{w} + \frac{w}{w'} + 4(\epsilon'\cdot\epsilon)^2 - 2\right] \tag{8.30}$$

Expression 8.30 is the Klein Nishina cross section for Compton scattering.

8.1 Addendum to Compton Scattering: Reduction to Classical (Thomson) result and the unpolarized cross-section

We had written down the following expression for the differential cross section per solid angle of the exiting photon

$$\frac{d\sigma}{d\Omega'} = e^4\frac{w'^2}{w^2} \times \frac{1}{4m^2} \times \left[\frac{w'}{w} + \frac{w}{w'} + 4(\epsilon'\cdot\epsilon)^2 - 2\right] \tag{8.31}$$

When we take the low energy limit of expression (8.31) [LEFTOVER: Why low energy: function of Thomson's apparatus] we obtain quite easily using $\frac{w'}{w}$ 1:

$$\frac{d\sigma}{d\Omega'} = \frac{e^4}{m^2}(\epsilon'\cdot\epsilon)^2 \tag{8.32}$$

The classical cross-section in equation (8.32) is obtained from classical considerations: An electromagnetic wave causes a charge to undergo oscillatory acceleration, and so radiates, the incoming wave has polarization ϵ and the outgoing ϵ'. This calculation is covered in several texts [REFERENCE NEEDED].

The cross section in (8.32) depends on each single possible polarization of each photon. Since in practice we may not observe either polarization we can sum over the possibilities for ϵ' and average over the possibilities for ϵ.

As mentioned earlier (REFERENCE NEEDED) we are free to pick the polarization 4-vectors to be purely spacelike. This will allow the aforementioned sums to be particularly simple.

(REFERENCE NEEDED, MODIFICATION OF PREVIOUS TEXT NEEDED, LEFTOVER: WHY EXACTLY? Lorenz gauge *and* $p_i = (m, 0)$ or is this independent of the choice of p_i?) For simplicities sake we choose to perform the polarization sum by considering the scattering in the plane formed by the spacelike components of k and k'. We allow $\vec{k}$ to be in the $\hat{z}$ direction and so are now free to pick:

$$\epsilon_1^\mu = (0, 1, 0, 0)$$

$$\epsilon_2^\mu = (0, 0, 1, 0)$$

Now since we have chosen the plane spanned by $\vec{k}$, $\vec{k}'$ we can choose $\epsilon_2'^\mu$ to point out of the plane as before thus deriving:

$$\epsilon_1'^\mu = (0, \cos\theta, 0, -\sin\theta)$$

$$\epsilon_2'^\mu = (0, 0, 1, 0)$$

where θ is the angle between the $\lambda = 1$ polarizations that are in the plane spanned by $\vec{k}'$, and $\vec{k}$.

with this information at hand we sum over the four terms:

$$\sum_{\lambda,\lambda'=}^{2} (\epsilon_\lambda \cdot \epsilon'_{\lambda'})^2 = \cos^2\theta + 0 + 1 + 0 = 2 - \sin^2\theta \tag{8.33}$$

Note that in equation (8.31) the term in Eqn. 8.32 is multiplied by 4 and then must be averaged ($\times\frac{1}{2}$) the contribution of this to the whole differential cross-section is thus:

$$e^4 \frac{w'^2}{w^2} \frac{1}{4m^2} \times (4 - 2\sin^2\theta) \tag{8.34}$$

combining this with the remainder of terms in Eqn 8.31 we obtain (after accounting for 4 sums and a factor of $\frac{1}{2}$ for the average):

$$\frac{d\sigma}{d\Omega'} = e^4 \frac{w'^2}{w^2} \frac{2}{4m^2} \times \left[\frac{w'}{w} + \frac{w}{w'} - 2 \right] + e^4 \frac{w'^2}{w^2} \times \frac{1}{4m^2} \times (4 - 2\sin^2\theta) = \frac{e^4 w'^2}{2m^2 w^2} \left[\frac{w'}{w} + \frac{w}{w'} - \sin^2\theta \right] \tag{8.35}$$

Eqn. 8.35 is thus the un-polarized cross-section for Compton scattering, where θ defines the angle between one of the two polarizations.

[LEFTOVER: Link this to scattering angle?]

8.2 Pair Annihilation into two photons $e^+e^- \to \gamma\gamma$

At the end of the chapter on propagators we discussed the following possible scenario described in the language of hole theory and using the properties of our propagator: A negative energy electron encounters a potential and is scattered into a negative energy state which of course propagates backward in time. We observe therefore a positron and an electron propagating into the interaction region. In the hole theory language we described a process by which an electron with $E_- > 0$ could transit to an $E_+ < 0$ unoccupied hole and give off $E_- + E_+$ in the form of radiation, in a description drawing an analogy with an atomic transition where an excited state transits to a lower energy state in an atom.

We need to correct this narrative somewhat in an atom momentum is absorbed by the atom itself. If we however denote the 4-momentum of a photon by k and the electron positron mome ntum by $p_- = (E_-, \vec{p}_-)$ $p_+ = (E_+, \vec{p}_+)$ going to the center of mass frame $\vec{p}_+ + \vec{p}_- = 0$, one can easily see that the requrement for conservation of 4-momentum cannot be satisfied:

$$k^2 = (p_- + p_+)^2 = (E_+ + E_-)^2 - (0)^2$$

$$k^2 = 0 = 4E^2 (E = E_+ \text{ or } E_-)$$

clearly (at least) a second photon is needed.

Thus:

$$(k_1 + k_2)^2 = 2k_1 \cdot k_2 = 4E^2$$

and it's obvious that this relation can be satisfied. We've just described the physically observed phenomenon of e^+e^- pair annihilation into a pair of photons. We can easily see that the process of pair annihilation is related to the Compton scattering by crossing symmetry. To write down the amplitude for their annihilation we need to make the now familiar substitutions:

(1) $u_i \to u_-$, $p_i \to p_-$. These two are notational changes only, and define the incoming electron.
(2) $\bar{u}_f \to \bar{v}_+$, $p_f \to -p_+$. The outgoing electron is replaced by an incoming positron.
(3) $k' \to k_1$. Change notation for one outgoing photon. $k \to -k_2$. incoming photon $\to$ outgoing photon.

With these modifications at hand we write down the amplitude S_{fi} for pair annihilation using Eqn. (REFERENCE NEEDED), the same gauge choices and straightforward familiar manipulation following the development in Compton scattering:

$$S_{fi} = -ie^2\sqrt{\frac{m^2}{E_+E_-\omega_1\omega_2}}\times\frac{2\pi}{V^2}(2\pi)^4\delta(k_1+k_2-p_+-p_-)\times\bar{V}_+\left[\frac{\not{\epsilon}_2\,\not{k}_1\,\not{\epsilon}_1}{2p_-\cdot k_1}+\frac{\not{\epsilon}_1\,\not{k}_2\,\not{\epsilon}_2}{2p_-\cdot k_2}\right]u_- \tag{8.36}$$

modulus squaring S_{fi} and averaging over the electron and positron spins using (REFERENCE NEEDED) $\sum_{S_+} V_+\bar{V}_+ = \frac{-(-\not{p}_++m)}{2m}$ we obtain easily:

$$\frac{|S_{fi}|^2}{T} = \frac{e^4m^2}{E_+E_-\omega_1\omega_2}\frac{(2\pi)^6}{V^3}\delta^4(k_1+k_2-p_+-p_-) \tag{8.37}$$

$$\times\frac{1}{4}\times(-1)Tr\left[\frac{(-\not{p}_++m)}{2m}\left(\frac{\not{\epsilon}_2\,\not{k}_1\,\not{\epsilon}_1}{2p_-\cdot k_1}+\frac{\not{\epsilon}_1\,\not{k}_2\,\not{\epsilon}_2}{2p_-\cdot k_2}\right)\times\frac{(\not{p}_-+m)}{2m}\left(\frac{\not{\epsilon}_1\,\not{k}_1\,\not{\epsilon}_2}{2p_-\cdot k_1}+\frac{\not{\epsilon}_2\,\not{k}_2\,\not{\epsilon}_1}{2p_-\cdot k_2}\right)\right]$$

at this point the manipulation of the trace term can simply be avoided and we can borrow the result after the manipulations between Eqns and Eqn. (REFERENCE NEEDED) used in Compton scattering. We note that the average over initial spin states now includes the positron as well, hence an extra factor of $\frac{1}{2}$ is incurred. Using $p_f \to -p_+$, $k' \to -k_2$, $k \to k_1$, $p_i \to p_-$ in expression (REFERENCE NEEDED) we obtain:

$$\frac{|S_{fi}|^2}{T} = \frac{e^4m^2(2\pi)^6}{E_+E_-\omega_1\omega_2}\times\frac{1}{V^3}\times\delta^4(k_1+k_2-p_+-p_-)$$

$$\times(-1)\times\frac{1}{64m^2}\times\left[-\frac{8p_-\cdot k_2}{p_-\cdot k_1}-\frac{8p_-\cdot k_1}{p_-\cdot k_2}+32(\epsilon_i\cdot\epsilon_2)^2-16\right]$$

We now take the initial electron to be stationary and after putting in the counting factors of $\frac{Vd^3\vec{k}_1}{(2\pi)^3}$ and $\frac{Vd^3\vec{k}_2}{(2\pi)^3}$ for the exiting photons we obtain the following expressions for the rate:

$$dR = \frac{e^4}{E_+ m\omega_1\omega_2 V}\times\delta^4(k_1+k_2-p_+-p_-)\times\frac{1}{8}\left[\frac{\omega_2}{\omega_1}+\frac{\omega_1}{\omega_2}-4(\epsilon_1\cdot\epsilon_2)^2+2\right]d^3\vec{k}_1 d^3\vec{k}_2 \tag{8.38}$$

To convert the rate in equation (8.38) to an differential cross section we must perform the usual divison by the flux of incoming particles for our particular choice of setup is a single incoming positron beam. The flux is the simply $\frac{|\vec{p}_+|}{VE_+}$ as we have seen in Chapter 5.

Doing this we obtain the following for the differential cross-section:

$$d\sigma = \frac{e^4\times\delta^4(k_1+k_2-p_+-p_-)}{m\omega_1\omega_2|\vec{p}_+|}\times\frac{1}{8}\left[\frac{\omega_2}{\omega_1}+\frac{\omega_1}{\omega_2}-4(\epsilon_1\cdot\epsilon_2)^2+12\right]\times d^3\vec{k}_1 d^3\vec{k}_2 \tag{8.39}$$

We will now integrate out one fo the final state photons variables and express the differential cross section in terms of the exit angle (solid) of the other. To do this we will as usual utilize the following relations first justified in (EQUATION REFERENCE NEEDED)

$$d^3\vec{k}_2 = \omega_2|\vec{k}_2|d\omega_2 d\Omega_2 = \omega_2^2 d\omega_2 d\Omega_2$$

$$\frac{d^3\vec{k}_1}{2\omega_1} = \theta(k_{10})\delta(k_1^2)\times d^4k_1$$

Plugging these into Eqn. 8.39 we obtain

$$\frac{d\sigma}{d\Omega_2} = \frac{e^4\omega_2}{m|\vec{p}_+|}\times\delta^4(k_1+k_2-p_+-p_-)\times\frac{1}{8}\left[\frac{\omega_2}{\omega_1}+\frac{\omega_1}{\omega_2}-4(\epsilon_1\cdot\epsilon_2)^2+2\right] \tag{8.40}$$

$$\times\theta(k_{10})\delta(k_1^2)\times 2\times d\omega_2\times d^4k_1$$

integrating Eqn. 8.40 over d^4k_1, one sets the values of the components of k_1 as required by the 4-dimensional δ function in Eqn. 8.40. One obtains then

$$k_1^2 = 0 = (p_+ + p_- - k_2)^2 = 2m^2 + 2mE_+ - 2m\omega_2 - 2E_+\omega_2 + 2|\vec{p}_+|\cos\theta \tag{8.41}$$

where θ is taken to be the angle between the trajectory of the incoming positron and *one* of the (indistinguishable) outgoing photons.

Using equation (8.41) we can write

$$\omega_2 = \frac{m(E_+ + m)}{m + E_+ - |\vec{p}_+|\cos\theta} \tag{8.42}$$

to integrate over $d\omega_2$ we will need to consider the fact that the remaining one dimensional δ function $\delta(k_1^2)$ now $\delta((p_+ + p_- - k_2)^2)$ is a function of ω_2 as the right hand side of Eqn. 8.41 shows, the integration will incurr a factor of $|\frac{\partial f(\omega_2)}{\partial\omega_2}|^{-1}$ where $f(\omega_2)$ is the right hand of (8.41). Using this and equation 8.42 we obtain:

$$\frac{d\sigma}{d\Omega_2} = \frac{e^4}{m|\vec{p}_+|} \times \frac{m(E+m)}{(m + E_+ - |\vec{p}_+|\cos\theta)^2} \times \frac{1}{8}\left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} - 4(\epsilon_1\cdot\epsilon_2)^2 + 2\right] \tag{8.43}$$

Once again by defining θ_{12} to be the angle between ϵ_1 and ϵ_2 in a manner following the Compton scattering discussion (REFERENCE NEEDED) and by noting that no average is to be taken here we arrive at the unpolarized differential cross-section:

$$\frac{d\sigma}{d\Omega_2} = \frac{e^4}{m|\vec{p}_+|} \frac{m(E+m)}{(m + E_+ - |\vec{p}_+|\cos\theta)^2} \times \frac{1}{8}\left[4\frac{\omega_2}{\omega_1} + 4\frac{\omega_1}{\omega_2} + 4\sin^2\theta_{12} - 8 + 8\right] \tag{8.44}$$

$$= \frac{e^4}{m|\vec{p}_+|} \frac{m(E+m)}{(m + E_+ - |\vec{p}_+|\cos\theta)^2} \times \frac{1}{2}\left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + \sin^2\theta_{12}\right]$$

LEFTOVERS: Limits of total cross-section REFERENCES to other texts needed.

Bibliography

1. J .D. Jackson(1998).*Classical Electrodynamics.* Third Edition, 1998. Hamilton Printing Company.
2. Arfken, G. *Mathematical Methods for Physicists*, 3rd ed. Orlando, FL: Academic Press, 1985.
3. Feynman, R. P., Weinberg, S. *Elementary Particles and the Laws of Physics*, Cambridge University Press.
4. Bevan B. Baker *The Mathematical Theory of Huygens' Principle*, Ams Chelsea Publishing, 2003.
5. Wolfgang K. H. Panovsky, Melba Phillips(2005).*Classical Electricity and Magnetism.* Dover Publications.
6. S. .T. Thornton, J. B. Marion(2004), *Classical Dynamics of Particles and Systems.* Brooks/Cole, 2004.
7. J. D. Bjorken and S. D. Drell (1964)*Relativistic Quantum Mechanics* McGraw-Hill.
8. W. Greiner(2000). *Relativistic Quantum Mechanics.* Third edition (2000), Springer.
9. W. Greiner, J. Rheinhardt(2002). *Quantum Electrodynamics.* Third edition (2002), Springer.
10. R. Parthasarathy(2010). *Relativistic Quantum Mechanics.* First edition (2010), Alpha Science International.
11. Gordon Baym, (1994). *Lectures on Quantum Mechanics.* Twentieth Printing (1994), Addison Wesley Publishing Company.
12. Douglas M. Gingrich *Practical Quantum Electrodynamics* First edition 2006. Taylor and Francis Group.
13. Feynman, R. P. 1949a. *The theory of positrons* Physical Review 76:749759.
14. Feynman, R. P. 1949b. *Space-time approach to quantum electrodynamics* Physical Review 76:769789.
15. Richard P. Feynman *Quantum Electrodynamics* Advanced Book Classics, Perseus Books, 1998.
16. D. Miller. *Relativistic Quantum Mechanics*

    `http://www.physics.gla.ac.uk/~dmiller/lectures/RQM_2008.pdf`

Index