1 Natural units

Almost all units in nature are derived. For instance, in the cgs system of units one chooses the unit of length to be cm, the unit of mass to be g and the unit of time to be s. All other units can be expressed in terms of these three. For instance, the unit of force can be found from Newton's law to be $\frac{ML}{T^2}$ where M, L and T represent the particular units chosen for mass, length and time respectively. Similarly, the unit of charge can be found from Coulomb's law to be $\frac{M^{\frac{1}{2}}L^{\frac{3}{2}}}{T}$. We see that sometimes odd-looking fractional powers appear which is probably one reason why people have invented new names for these units. Anyway, the choice of units which are regarded as fundamental is by no means unique. For instance, instead of choosing M, L and T, we can choose as fundamental, the unit of energy E, the unit of velocity V and the unit of "action" A. The last unit is somewhat unusual. It is the unit that the action functional in mechanics or quantum mechanics carries. Since the action is $\int dt\ (K - V)$ we see that the unit is energy·time, the same as the constant $\hbar$ carries. (It would also be possible to view A as the unit of angular momentum.) If we choose the units for velocity such that the speed of light, $c = 1$ and the unit of action such that $\hbar = 1$, then we have something that is called "natural units". It only remains to choose a unit for energy, which we usually will choose to be eV.

To be able to convert between these two systems of units, we express the new units in terms of the old ones. It is not difficult to find that

$$\left\{\begin{array}{rcl} V & = & \frac{L}{T} \\ E & = & \frac{ML^2}{T^2} \\ A & = & \frac{ML^2}{T} \end{array}\right. \quad \text{or} \quad \left\{\begin{array}{rcl} M & = & \frac{E}{V^2} \\ T & = & \frac{A}{E} \\ L & = & \frac{AV}{E} \end{array}\right. \tag{1}$$

Using these relations we can make a small table

Quantity	CGS	Natural
Force	$\frac{ML}{T^2}$	$\frac{E^2}{AV}$
Charge	$\frac{M^{\frac{1}{2}}L^{\frac{3}{2}}}{T}$	$(AV)^{\frac{1}{2}}$
Magnetic moment	$\frac{M^{\frac{1}{2}}L^{\frac{5}{2}}}{T}$	$\frac{(AV)^{\frac{3}{2}}}{E}$

From the table we see that for instance charge is dimensionless in natural units (it is given in eV^0). Force has dimensions eV^2 while length, time and magnetic moment all have the same dimension eV^{-1}. If you are given a number in natural units and you want to change it into CGS units you just have to insert the proper powers of $\hbar$ and c to restore the dimensions. Here is an example: the charge of the electron in natural units is

$$e_{\mathrm{n}} = 8.543 \cdot 10^{-2}\ eV^{0} \tag{2}$$

Converting this to CGS we have to consult the table to see that we need to multiply this with $\sqrt{\hbar c}$ to restore the dimensions. Note here that the units you choose for $\hbar$ and c will also give you the units of the final result. Since we are interested in CGS units we have to give $\hbar$ and c in cm, g and s instead of the more usual m, kg and s. For reference I give them here

$$\begin{aligned} \hbar &= 1.05459 \cdot 10^{-27}\ \frac{\mathrm{g\ cm}^2}{\mathrm{s}} \\ c &= 2.9979 \cdot 10^{10}\ \frac{\mathrm{cm}}{\mathrm{s}} \end{aligned} \tag{3}$$

Multiplying together gives

$$e_{\mathrm{cgs}} = e_{\mathrm{n}}\sqrt{\hbar c} = 4.803 \cdot 10^{-10}\ \mathrm{esu} \tag{4}$$

Another example is is the expression for the Bohr radius in natural units

$$(a_0)_{\mathrm{n}} = \frac{1}{m_{\mathrm{n}} e_{\mathrm{n}}^2}\ eV^{-1} \tag{5}$$

To transform this to an expression of dimension length we multiply with $\hbar c$ to get

$$(a_0)_{\mathrm{cgs}} = \frac{\hbar c}{m_{\mathrm{n}} e_{\mathrm{n}}^2} \tag{6}$$

The formula still depends on the values of e and m in natural units though. We can convert these also by using the relation between the charges derived above and the relation between the masses as $m_{\mathrm{n}} = m_{\mathrm{cgs}} c^2$. This gives

$$(a_0)_{\mathrm{cgs}} = \frac{\hbar^2 c^2}{m_{\mathrm{cgs}} c^2 e_{\mathrm{cgs}}^2} = \frac{\hbar^2}{m_{\mathrm{cgs}} e_{\mathrm{cgs}}^2} \tag{7}$$

Comment 1: The definition and even the *unit* of charge differs between different systems of units. This comes about because there are different conventions about how to write Coulomb's law. We have

$$e_{\mathrm{cgs}} = \frac{e_{\mathrm{hl}}}{\sqrt{4\pi}} = \frac{e_{\mathrm{SI}}}{\sqrt{4\pi\varepsilon_0}} \tag{8}$$

$e_{\rm hl}$ represents the charge in the Heaviside-Lorentz system of units. It clearly has the same dimension as the CGS charge. However, in the SI system there is an additional constant ε_0 which has the dimension of inverse velocity squared. Charge in the SI system therefore has a different dimension than in CGS.

Comment 2: There exists other systems of units with only one basic unit. For instance, in relativistic gravitational physics it is often advantageous to choose a system of units where the speed of light $c = 1$ and Newton's gravitational constant $G = 1$. The remaining unit is the unit of length which can be chosen arbitrarily (light-year, m, cm etc.). This system of units is called *geometrical units*.

2 The Dirac equation

There is a curious way to "derive" the Schrödinger equation. Namely, take the relation for the energy in classical physics

$$E = \frac{p^2}{2m} + V. \tag{9}$$

One gets the Schrödinger equation by making the replacement

$$\begin{aligned} E &\rightarrow i\hbar\frac{\partial}{\partial t}, \\ p_i &\rightarrow -i\hbar\frac{\partial}{\partial x^i}, \end{aligned} \tag{10}$$

and then letting the relation (9) "act" on a wavefunction one gets

$$i\hbar\frac{\partial}{\partial t}\psi = \left(-\frac{\hbar^2}{2m}\partial_x^2 + V\right)\psi. \tag{11}$$

This derivation inspired many people to try to derive a relativistic analog of the Schrödinger equation by starting with the relativistic energy relation $E^2 = \mathbf{p}^2c^2 + m^2c^4$ (or $E^2 = \mathbf{p}^2 + m^2$ in natural units) instead of starting with (9). Making the same substitution (10) as before we get a *relativistic* wave equation

$$-\partial_t^2\phi = -\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)\phi + m^2\phi. \tag{12}$$

This can be written in a more relativistic fashion by introducing a metric $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ as

$$g^{\mu\nu}\partial_\mu\partial_\nu\phi + m^2\phi = 0, \tag{13}$$

an equation which is known as the Klein-Gordon equation.

To find out more about its properties, we now go on to find solutions to the Klein-Gordon equation. For instance, there is a complete set of plane-wave solutions as we will now show. First we make the ansatz $\phi = e^{-ik_\mu x^\mu}$. Acting on this with a four-derivative ∂_μ gives us

$$\partial_\mu e^{-ik_\nu x^\nu} = -ik_\mu e^{-ik_\nu x^\nu}. \tag{14}$$

Using this result twice we may insert the ansatz into the Klein-Gordon equation to get

$$\partial_\mu\partial^\mu\phi + m^2\phi = (-k_\mu k^\mu + m^2)\phi. \tag{15}$$

We see that for ϕ to be a solution to the Klein-Gordon equation we need the four momentum k_μ to satisfy the relation

$$k_\mu k^\mu = m^2, \tag{16}$$

and rewriting the four momentum k_μ in terms of its components $k^\mu = (E, \mathbf{k})$ where $\mathbf{k}$ is the ordinary three momentum, we recover the relativistic energy relation $E^2 = \mathbf{k}^2 + m^2$. Let us recapitulate; the Klein-Gordon equation has a complete set of plane wave solutions $\phi(x) = e^{ik\cdot x}$ where the four momentum has to satisfy the relativistic energy condition $k \cdot k = m^2$. Any solution can then be written as a linear combination of these plane waves. There is however a funny new feature of these solutions. If the four vector $k^\mu = (E, \mathbf{k})$ gives a solution, then the four vector $k^\mu = (-E, \mathbf{k})$ with *negative* energy is also a solution! Thus, for every solution with positive energy, there is a solution with negative energy which seems physically unacceptable since it would lead to an unstable theory (there would be no state with lowest energy = vacuum state).

Dirac identified the root of this problem in the fact that the Klein-Gordon equation is *quadratic* in time derivative whereas the Schrödinger equation is linear. He tried to get around this by making an equation which would be linear in time derivatives. To achieve this he used some interesting properties of the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{17}$$

which fulfill the relation $\sigma^i\sigma^k = i\epsilon^{ikl}\sigma^l + \delta^{ik}\mathbf{1}$. This made it possible for Dirac to write

$$k_\mu k^\mu = E^2 - \mathbf{k}^i\mathbf{k}^i = \left(E\mathbf{1} - \mathbf{k}^i\sigma^i\right)\left(E\mathbf{1} + \mathbf{k}^l\sigma^l\right). \tag{18}$$

That is, by writing the equation in terms of two by two matrices, he was able to split it into factors linear in energy. The price he had to pay was that the wave function now become two dimension column vectors (or spinors as they are more commonly known). Thus our second try for a relativistic wave equation looks like this

$$\left(\mathbf{1}i\partial_t - \sigma^i i\partial_i\right)\left(\mathbf{1}i\partial_t + \sigma^l i\partial_l\right)\phi_A = m^2\phi_A, \tag{19}$$

where $\phi_A = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$ is a two dimensional column vector. By introducing a second two dimensional column vector

$$m\phi_B = \left(\mathbf{1}i\partial_t + \sigma^l i\partial_l\right)\phi_A, \tag{20}$$

we can write an equation (well, really a system of equations) which is linear in time derivatives

$$\begin{aligned} m\phi_B &= \left(\mathbf{1}i\partial_t + \sigma^l i\partial_l\right)\phi_A, \\ m\phi_A &= \left(\mathbf{1}i\partial_t - \sigma^i i\partial_i\right)\phi_B. \end{aligned} \tag{21}$$

For purely conventional reasons one often redefines the column vectors as $\phi_\pm = \phi_A \pm \phi_B$ which makes it possible to write the above equation as

$$\begin{aligned} m\phi_+ &= \mathbf{1}i\partial_t\phi_+ + \sigma^l i\partial_l\phi_- \\ m\phi_- &= -\mathbf{1}i\partial_t\phi_- - \sigma^l i\partial_l\phi_+, \end{aligned} \tag{22}$$

or, defining a four component column vector $\psi = \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}$ and four by four matrices

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}; \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{23}$$

we may write the resulting equations in a very compact form as

$$\gamma^\mu i\partial_\mu\psi = m\psi. \tag{24}$$

Notice that this is a matrix equation (it is really four equations written in a very nice and compact form using matrices) and that it is linear in time derivatives which is exactly what Dirac wanted to achieve. This equation is known as the Dirac equation. To make the comparison to the ordinary Schrödinger equation more prominent, we can rewrite it as

$$\gamma^0 i\partial_t \psi = -\gamma^l i\partial_l \psi + m\psi, \tag{25}$$

and using that $\gamma^0\gamma^0 = 1$ we find

$$i\partial_t \psi = -\gamma^0\gamma^l i\partial_l \psi + m\gamma^0\psi. \tag{26}$$

We thus see that the Hamiltonian operator that we get from the Dirac equation is $H = -\gamma^0\gamma^l i\partial_l + m\gamma^0$.

Again, to get a feeling for the physics we can try to solve the equation. Since the wavefunction is a four component column vector we make the ansatz for a plane wave

$$\psi = u(p)e^{-ip\cdot x}, \tag{27}$$

where $u(p)$ is a four component column vector possibly dependent on $\mathbf{p}$. Inserting this into the Dirac equation we get

$$(i\gamma^\mu\partial_\mu - m)\psi = (\gamma^\mu p_\mu - m)u(p)e^{-ip\cdot x}, \tag{28}$$

so we see that for this to be a solution of the Dirac equation we need the four column vector u to satisfy the matrix equation

$$(\gamma^\mu p_\mu - m)u(p) = 0. \tag{29}$$

Using the expressions for the gamma matrices found earlier we can rewrite this in an even more explicit form

$$\begin{pmatrix} E-m & 0 & -p_3 & -p_- \\ 0 & E-m & -p_+ & p_3 \\ p_3 & p_- & -E-m & 0 \\ p_+ & -p_3 & 0 & -E-m \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0, \tag{30}$$

where we have defined the complex combinations $p_\pm = p_1 \pm ip_2$. This equation has four independent solutions. We will find one of them, but I recommend that you similarly try to find the other three. Actually, for this equation to be solvable we need the determinant of the matrix to be zero. We can easily evaluate it to be $(E^2 - \mathbf{p}^2 - m^2)^2$ so we see that a necessary condition for this equation to have solutions is that the "old" relativistic energy condition is satisfied. Unfortunately this means that we did not get rid of the solutions with negative energy. Therefore we first need to assume that the condition holds, then we can go on and try to find a solution. To make it a little bit simpler, let us first try in the case where $\mathbf{p} = 0$. Then the equation looks like

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2m & 0 \\ 0 & 0 & 0 & -2m \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0, \tag{31}$$

for the case of positive energy, i.e. when $E = +m$ and in the case where the energy is negative, i.e. when $E = -m$, it looks like

$$\begin{pmatrix} -2m & 0 & 0 & 0 \\ 0 & -2m & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \tag{32}$$

In the positive energy case we have the two independent solutions

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{33}$$

and in the negative energy case the solutions look like

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{34}$$

Turning on the three momentum $\mathbf{p}$ we have to solve the full equations (30) but we can expect that the solutions should not differ too much from the zero $\mathbf{p}$ solutions, at least when $\mathbf{p}$ is small. Then we should be able to find a solution of the form

$$\begin{pmatrix} 1 \\ 0 \\ a \\ b \end{pmatrix}, \tag{35}$$

where a and b are small of order $\mathbf{p}$ (or possibly smaller). Inserting this ansatz into the equation immediately gives us that $a = \frac{p_3}{E+m}$ and $b = \frac{p_+}{E+m}$. For reasons to be explained later we choose the normalization to be $u^\dagger u = 2E$ which leads to the final answer

$$\psi = \sqrt{E+m} \begin{pmatrix} 1 \\ 0 \\ \frac{p_3}{E+m} \\ \frac{p_+}{E+m} \end{pmatrix} e^{-ik\cdot x}. \tag{36}$$

3 The non-relativistic limit of the Dirac equation

One check that one should always do is to see how the new physics one is investigating reduces in known situations. In the case at hand this means that we should try to see how the physics of the Dirac equation looks in a non-relativistic situation. To do this, let us have a look at it in the form given in (22) but in momentum space. The equation looks like

$$\begin{aligned} (E-m)\phi_+ &= \sigma^l p_l \phi_-, \\ (E+m)\phi_- &= \sigma^l p_l \phi_+. \end{aligned} \tag{37}$$

The non-relativistic limit means the limit where $\mathbf{p} \ll m$. This in turn implies that $E = \sqrt{\mathbf{p}^2 + m^2} = m + \frac{\mathbf{p}^2}{2m} + \ldots = m + E^{(NR)}$ where $E^{(NR)}$ is the non-relativistic energy. This immediately tells us that the quantity $E^{(NR)} = E - m$ is small (of order $m\left(\frac{\mathbf{p}}{m}\right)^2$ or $\frac{v^2}{c^2}mc^2$ in ordinary units) while the quantity $E+m$ is large (of order m). A look at the equations now tell us that ϕ_+ is of order one while ϕ_- is of order $\frac{\mathbf{p}}{m}$ so it goes to zero in the non-relativistic limit. We can now solve for the "small" component ϕ_- to get an equation for ϕ_+ only since ϕ_+ is what is left in the non-relativistic limit. Solving for ϕ_- gives us

$$\phi_- = \frac{1}{2m + E^{(NR)}} \mathbf{p}\cdot\sigma\phi_+, \tag{38}$$

which, when inserted back into the equation gives us

$$E^{(NR)}\phi_+ = \mathbf{p}\cdot\sigma \frac{1}{2m + E^{(NR)}} \mathbf{p}\cdot\sigma\phi_+. \tag{39}$$

In the non-relativistic limit $m \gg E^{(NR)}$ so we can expand the denominator to get

$$E^{(NR)}\phi_+ = \mathbf{p}\cdot\sigma\frac{1}{2m}(1 - \frac{E^{(NR)}}{2m} + \ldots)\mathbf{p}\cdot\sigma\phi_+, \tag{40}$$

and to lowest order we get back the non-relativistic Schrödinger equation

$$E^{(NR)}\phi_+ = \frac{\mathbf{p^2}}{2m}\phi_+. \tag{41}$$

This is maybe not a very exciting result but it is gratifying to see that we get the correct non-relativistic limit of our equation.

A slightly more interesting result we get if we include a potential from an external electromagnetic field. This is done in a relativistically covariant fashion in the Dirac equation, introducing the relativistic electromagnetic vector potential $A^\mu = (\varphi, \mathbf{A})$, by replacing $i\partial_\mu \to i\partial_\mu - eA_\mu$. This changes the Dirac equation to

$$\gamma^\mu \left(i\partial_\mu - eA_\mu\right)\psi = m\psi, \tag{42}$$

or, if we Fourier transform as

$$\gamma^\mu \left(p_\mu - eA_\mu\right)u(p) = mu(p). \tag{43}$$

When rewriting this in terms of the large and small components we get

$$\begin{aligned}(E - e\varphi - m)u_+ &= \sigma\cdot(\mathbf{p} - e\mathbf{A})u_-,\\ (E - e\varphi + m)u_- &= \sigma\cdot(\mathbf{p} - e\mathbf{A})u_+,\end{aligned} \tag{44}$$

and solving for the small component we get

$$\begin{aligned}u_- &= \frac{1}{2m + E^{(NR)} - e\varphi}\sigma\cdot(\mathbf{p} - e\mathbf{A})u_+,\\ E^{(NR)}u_+ &= \left(e\varphi + \sigma\cdot(\mathbf{p} - e\mathbf{A})\frac{1}{2m + E^{(NR)} - e\varphi}\sigma\cdot(\mathbf{p} - e\mathbf{A})\right)u_+.\end{aligned} \tag{45}$$

Notice that we have to be careful in which order we write things since φ and $\mathbf{A}$ depend on $\mathbf{x}$ and thus do not commute with $\mathbf{p}$. Using the same approximations as before we get and equation for u_+

$$E^{(NR)}u_+ = \left(e\varphi + \frac{1}{2m}\sigma\cdot(\mathbf{p} - e\mathbf{A})\sigma\cdot(\mathbf{p} - e\mathbf{A})\right)u_+. \tag{46}$$

To evaluate this we again need to use the properties of the Pauli matrices to be able to write

$$
\begin{aligned}
\sigma\cdot(\mathbf{p}-e\mathbf{A})\sigma\cdot(\mathbf{p}-e\mathbf{A}) = \sigma^i\sigma^k(\mathbf{p}-e\mathbf{A})_i(\mathbf{p}-e\mathbf{A})_k = \\
(\delta^{ik}+i\epsilon^{ikl}\sigma^l)(\mathbf{p}-e\mathbf{A})_i(\mathbf{p}-e\mathbf{A})_k = \\
(\mathbf{p}-e\mathbf{A})\cdot(\mathbf{p}-e\mathbf{A}) + i\sigma\cdot(\mathbf{p}-e\mathbf{A})\times(\mathbf{p}-e\mathbf{A}).
\end{aligned}
\tag{47}
$$

The cross product can be evaluated as

$$
\begin{aligned}
\epsilon^{lik}(\mathbf{p}-e\mathbf{A})_i(\mathbf{p}-e\mathbf{A})_k = \frac{1}{2}\epsilon^{lik}\left[(\mathbf{p}-e\mathbf{A})_i,(\mathbf{p}-e\mathbf{A})_k\right] = \\
\frac{1}{2}\epsilon^{lik}\left(-e\left[\mathbf{p}_i,\mathbf{A}_k\right]-e\left[\mathbf{A}_i,\mathbf{p}_k\right]\right) = -\epsilon^{lik}e\left[\mathbf{p}_i,\mathbf{A}_k\right] = \qquad (48)\\
ie\epsilon^{lik}\partial_i\mathbf{A}_k = ie\mathbf{B}^l, \qquad (49)
\end{aligned}
$$

which gives us the non-relativistic equation

$$
E^{(NR)}u_+ = \left(\frac{(\mathbf{p}-e\mathbf{A})^2}{2m} + e\varphi - \frac{e}{2m}\sigma\cdot\mathbf{B}\right)u_+. \tag{50}
$$

This is exactly the Schrödinger equation for a non-relativistic spin half particle with an intrinsic magnetic moment $\mu = \frac{e}{m}\mathbf{s}$ where $\mathbf{s} = \frac{\sigma}{2}$ is the spin operator. This is a very interesting result. We see that *only* from the requirement that the theory should be relativistically invariant, we find that particles carry an intrinsic magnetic moment. This is not something that we can turn off or change in any way. It is fundamentally built into the theory and comes from the relativistic invariance. Furthermore, it cannot be understood in any classical sense as "something charged going around in circles". In fact, you can easily verify by yourself that if we have some charged particle moving in a circle of radius R it produces a magnetic moment which is $\mu = \frac{e}{2m}\mathbf{L}$ and what we get out of our equation is twice this value. We say that the electron has a gyromagnetic ratio of 2. In fact this is not completely true and this value receives quantum corrections which can be computed with great accuracy (moreover, you will in principle be able to do it yourself using what you learn in this course).

One can go on and keep higher order corrections to this result. This will result in extra terms in the Hamiltonian. The calculation is slightly more involved since now it will not be justified to neglect ϕ_- any more. Anyway, it is still possible to write a non-relativistic Hamiltonian for a two component spinor. If one puts $\mathbf{A} = 0$ (no magnetic field) the Schrödinger equation becomes

$$\left(\frac{\mathbf{p^2}}{2m} + e\varphi - \frac{\mathbf{p^4}}{8m^3} - \frac{e\sigma\cdot(\mathbf{E}\times\mathbf{p})}{4m^2} - \frac{e}{8m^2}\nabla\cdot\mathbf{E}\right)\psi = E^{(NR)}\psi. \tag{51}$$

The first two terms are the lowest order terms which we have already derived (remember that we put $\mathbf{A} = 0$). The next three terms are higher order corrections. If we for instance apply this Hamiltonian to the hydrogen atom they will give small corrections to the spectrum (known as *fine structure*). The third term is simply the first non-trivial correction to the non-relativistic energy (from expanding $\sqrt{\mathbf{p}^2 + m^2} - m$). The fourth term is called the Thomas term and it has the interpretation as an interaction between the spin of the electron and the effective magnetic field it sees when moving through the electric field. It can be rewritten as a spin-orbit interaction (proportional to $\mathbf{S}\cdot\mathbf{L}$). The last term is known as the Darwin term. It represents an interaction with the charge density that produces the electric field. In the hydrogen atom it gives a shift in energy of the s-states. There is also something called *hyperfine structure* of the hydrogen spectrum. It comes from the interaction of the magnetic moments of the proton and the electron but is a much smaller effect than is the fine structure.

I would like to point out once again that all these terms one get automatically from the Dirac equation when going to the non-relativistic limit. There are no additional assumptions involved. Quite a nice little equation!

4 Transformation properties of the Dirac equation

You are familiar with how covariant and contravariant vectors transform when we change coordinate systems (we also say "when we do Lorentz rotations" or "boost" the coordinate system) in special relativity. The typical contravariant vector is the coordinate vector x^μ itself. When we do a Lorentz boost it transforms into $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ where, if we for instance boost to a coordinate system which is moving with speed v in the x direction we have

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & -\frac{v}{\sqrt{1-v^2}} & 0 & 0 \\ -\frac{v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{52}$$

We may define the matrix $\Lambda_\mu{}^\nu = g_{\mu\rho}\Lambda^\rho{}_\sigma g^{\sigma\nu}$ and we can check that $\Lambda^\rho{}_\mu \Lambda_\rho{}^\nu = \delta^\nu_\mu$. All covariant quantities (for example the momentum vector p_μ or a vector field A_μ or the ordinary derivative operator ∂_μ transform as $A'_\mu = \Lambda_\mu{}^\nu A_\nu$. Therefore the scalar product is invariant $x'^\mu p'_\mu = x^\mu p_\mu$. Using this information it is easy to see that for a scalar field ϕ (a scalar field is defined by the property that it does not transform at all under Lorentz transformations) the Klein-Gordon equation is invariant under Lorentz transformations

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0 \tag{53}$$

A spinor is not invariant under Lorentz transformations but transforms as $\psi'_a = S_{ab}\psi_b$ for some matrix S which we will not need the exact form of. The Dirac equation itself transforms as

$$i\partial\!\!\!/\psi - m\psi = 0 \to i\gamma^\mu \Lambda_\mu{}^\sigma \partial_\sigma (S\psi) - mS\psi = 0 \tag{54}$$

or

$$iS^{-1}\gamma^\mu S \Lambda_\mu{}^\sigma \partial_\sigma \psi - m\psi = 0 \tag{55}$$

We see that for the Dirac equation to be invariant we need that

$$S^{-1}\gamma^\mu S = \Lambda^\mu{}_\sigma \gamma^\sigma \tag{56}$$

Taking the hermitian conjugate of this equation and using that we know from the explicit representation of the gamma matrices that $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$ we get

$$\gamma^0 S^\dagger \gamma^0 = S^{-1} \tag{57}$$

Having this formula we may investigate how for instance $\psi^\dagger$ transforms under Lorentz transformations. We get

$$\psi'^\dagger = \psi^\dagger \gamma^0 S^{-1} \gamma^0 \tag{58}$$

So the hermitian conjugate does not transform as the inverse of the original object. However, if we check how $\bar\psi \equiv \psi^\dagger\gamma^0$ transforms we find

$$\bar\psi' = \bar\psi S^{-1} \tag{59}$$

which is indeed “nicer” since we can form objects with simple Lorentz transformation properties from it, for instance

$$\begin{aligned} \bar\psi'\psi' &= \bar\psi\psi \;\; \text{(scalar)} \\ \bar\psi'\gamma^\mu\psi' &= \bar\psi S^{-1}\gamma^\mu S\psi = \Lambda^\mu{}_\nu \bar\psi\gamma^\nu\psi \;\; \text{(vector)} \end{aligned} \tag{60}$$

5 Second quantization

To be able to describe quantum systems where the number of particles is able to change (for instance, an electron and a positron annihilates into two photons) we use a formalism called "second quantization". Notice that the name second quantization is rather badly chosen since it is not a question about "quantizing again". It is simply yet an alternative formalism for describing the states we have in the quantum world. It is not only used in relativistic quantum mechanics, but also in for instance solid state physics or anywhere were our quantum system consists of many types of particles which can also change into each other.

As a technical detail to simplify computations, let us imagine that our universe is a box with side length L. Then the universe has finite volume $V = L^3$ and if we impose periodical boundary conditions the allowed momenta form a countable set. In this universe the allowed momenta can be written as

$$\mathbf{k} = \frac{2\pi}{L}\left(n_1, n_2, n_3\right), \tag{61}$$

for any integers n_1, n_2, n_3. In the end of each calculation we may let $L \to \infty$ (if we have done the calculation correctly, nothing should depend on L).

Imagining that we have ordered the allowed momenta in some particular way, we may write them as $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_i, \ldots$. This gives us the possibility to write an arbitrary state of the system as

$$|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_\mathrm{i}}, \ldots\rangle, \tag{62}$$

which we interpret as meaning: there are $n_{\mathbf{k}_1}$ particles with momenta $\mathbf{k}_1$ (that means plane waves), there are $n_{\mathbf{k}_2}$ particles with momenta $\mathbf{k}_2$ and so on. These states in fact form a complete basis so any state can be written as a linear combination of these basis states. To be able to write down how operators act on these states we consider the "basic" operators $a_\mathbf{k}$ and $a_\mathbf{k}^\dagger$ satisfying the commutation relations $\left[a_\mathbf{p}, a_\mathbf{k}^\dagger\right] = \delta_{\mathbf{p},\mathbf{k}}$. That is, if $\mathbf{p}$ and $\mathbf{k}$ are different then $a_\mathbf{p}$ and $a_\mathbf{k}^\dagger$ commute but if they are the same they satisfy the usual harmonic oscillator algebra. Then, remembering the harmonic oscillator, we have that

$$\begin{aligned} a_{\mathbf{k}_i}|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_\mathrm{i}}, \ldots\rangle, &= \sqrt{n_{\mathbf{k}_i}}|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_\mathrm{i}} - 1, \ldots\rangle, \\ a_{\mathbf{k}_i}^\dagger|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_\mathrm{i}}, \ldots\rangle, &= \sqrt{n_{\mathbf{k}_i} + 1}|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_\mathrm{i}} + 1, \ldots\rangle. \end{aligned} \tag{63}$$

Thus, the $a, a^\dagger$ operators describe the basic operations which take us between different states. For instance, a process where we a particle with momentum $\mathbf{k}_1$ is scattered into a particle with momentum $\mathbf{k}_2$ would be accomplished by the operator $a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_1}$ acting on the state $|1_{\mathbf{k}_1}, 0, \ldots\rangle$. Explicitly, using (63), we would have

$$a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_1} |1_{\mathbf{k}_1}, 0, \ldots\rangle = a^\dagger_{\mathbf{k}_2} |0, 0, \ldots\rangle = |0, 1_{\mathbf{k}_2}, \ldots\rangle. \tag{64}$$

As in the harmonic oscillator, any state in the basis can be constructed by acting with the $a^\dagger$ operators on the vacuum

$$|\mathrm{n}_{\mathbf{k}_1}, \mathrm{n}_{\mathbf{k}_2}, \ldots, \mathrm{n}_{\mathbf{k}_i}, \ldots\rangle = \frac{\left(a^\dagger_{\mathbf{k}_1}\right)^{n_{\mathbf{k}_1}}}{\sqrt{n_{\mathbf{k}_1}!}} \frac{\left(a^\dagger_{\mathbf{k}_2}\right)^{n_{\mathbf{k}_2}}}{\sqrt{n_{\mathbf{k}_2}!}} \cdots \frac{\left(a^\dagger_{\mathbf{k}_i}\right)^{n_{\mathbf{k}_i}}}{\sqrt{n_{\mathbf{k}_i}!}} \cdots |0\rangle. \tag{65}$$

Let us now look at the coordinate representation of these states. Since we know the total number of particles in each state we know how many coordinates we need, i.e. one particle have three coordinates $\mathbf{x}$, two particles have six coordinates $\mathbf{x}_1, \mathbf{x}_2$ etc. Therefore we have

$$\begin{aligned} \langle \mathbf{x} | 1_{\mathbf{k}} \rangle &= \phi_{\mathbf{k}}(\mathbf{x}) \\ \langle \mathbf{x}_1, \mathbf{x}_2 | 1_{\mathbf{k}_1}, 1_{\mathbf{k}_2} \rangle &= \phi_{\mathbf{k}_1}(\mathbf{x}_1) \phi_{\mathbf{k}_2}(\mathbf{x}_2), \end{aligned} \tag{66}$$

where we have denoted the coordinate representation of the state with momentum $\mathbf{k}$ as $\phi_{\mathbf{k}} = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{x}}$. The factor $\frac{1}{\sqrt{V}}$ is a normalization factor.

Now consider the operator

$$\phi(\mathbf{x}) = \sum_{\mathbf{k}} a_{\mathbf{k}} \phi_{\mathbf{k}}(\mathbf{x}). \tag{67}$$

When its hermitian conjugate acts on the vacuum, it creates a state

$$\phi^\dagger(\mathbf{x}_0)|0\rangle = \sum_{\mathbf{k}} \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{x}_0} a^\dagger_{\mathbf{k}} |0\rangle = \sum_{\mathbf{k}} |1_{\mathbf{k}}\rangle \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{x}_0}. \tag{68}$$

In the coordinate representation, the resulting state looks like

$$\langle \mathbf{x} | \phi^\dagger(\mathbf{x}_0) | 0 \rangle = \sum_{\mathbf{k}} \langle \mathbf{x} | 1_{\mathbf{k}} \rangle \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{x}_0} = \frac{1}{V} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)}. \tag{69}$$

The final expression may seem a little bit strange but it is really a delta function. If $\mathbf{x} \neq \mathbf{x}_0$ the exponential oscillates for each $\mathbf{k}$ and on average it is zero. For $\mathbf{x} = \mathbf{x}_0$ however, all exponentials are 1 independent of $\mathbf{k}$ so the sum diverges. The integral over $\mathbf{x}$ of this function gives zero for all terms with $\mathbf{k} \neq 0$ and V for the $\mathbf{k} = 0$ term. The factor $\frac{1}{V}$ ensures that the final result of the integration is 1. Thus we see that the operator $\phi^\dagger(\mathbf{x}_0)$ creates a wavefunction which in coordinate representation is a delta function located at $\mathbf{x}_0$ or in other words, the operator $\phi^\dagger(\mathbf{x}_0)$ creates a particle completely localized at $\mathbf{x}_0$ complementary to $a^\dagger_{\mathbf{k}}$ which creates a particle with fixed momentum $\mathbf{k}$. In the same way, the operator $\phi(\mathbf{x}_0)$ annihilates a particle located at $\mathbf{x}_0$.

Now consider any operator on the full system that can be thought of as being composed of operations on the single particles. In formulas we would write

$$\mathcal{O} = \sum_a \mathcal{O}_a, \tag{70}$$

where the sum is over all particles in the system. This is a very general expression. Many operators of physical interest are of this type. For instance, the total energy is the sum of the energy of the single particles. The total momentum is the sum of the momenta of the single particles etc.

The operator $\mathcal{O}_a$ is a "normal" one particle operator. Its action on a one particle state can be expanded into a linear combination of one particle states

$$\mathcal{O}_x \phi_a(\mathbf{x}) = \sum_b \phi_b(\mathbf{x}) f_{ba}, \tag{71}$$

where, as usual

$$f_{ba} = \int d^3\mathbf{x}\ \phi^*_b(\mathbf{x}) \mathcal{O}_x \phi_a(\mathbf{x}). \tag{72}$$

Thus the action of each of the one particle operators in (70) can be seen as a reshuffling of the particles. The total number of particles is not changed, but they are moved between different states.

If we want to write how the operator (70) acts on the basis (62) we know that since it does not change the total number of particles but rather shuffles them around, it has to be written as a linear combination of the operators $a^\dagger_a a_b$ since this operator first annihilates a particle in state b but then immediately creates a particle in state a. Explicitly we write

$$\mathcal{O} = \sum_{ab} f_{ab} a^\dagger_a a_b, \tag{73}$$

and you should check that the coefficients f_{ab} are really the same as in (72) by for instance check how $\mathcal{O}$ acts on one particle states. Thus we may write

$$\mathcal{O} = \sum_{ab} \int d^3\mathbf{x}\ \phi_b^*(\mathbf{x})\mathcal{O}_x\phi_a(\mathbf{x})a_a^\dagger a_b = \tag{74}$$
$$\int d^3\mathbf{x}\ \left(\sum_a a_a\phi_a(\mathbf{x})\right)^\dagger \mathcal{O}_x \left(\sum_b a_b\phi_b(\mathbf{x})\right) = \int d^3\mathbf{x}\ \phi^\dagger(\mathbf{x})\mathcal{O}_x\phi(\mathbf{x}),$$

where we have used the operator $\phi(\mathbf{x})$ defined in (67). From this we read of the procedure for writing operators (this works only for operators that can be thought of as being composed of operations on the single particles) in the second quantized formalism. Take the one particle operator (here written $\mathcal{O}_x$) and compute what *looks like* an expectation value but instead of a wave function we use the *operator* $\phi(\mathbf{x})$. Since we know what the operator $\phi(\mathbf{x})$ does we know how to interpret this expression intuitively. First the operator ϕ annihilates a particle located at $\mathbf{x}$ (if there is a particle there, otherwise the result is zero), then the operator $\mathcal{O}_x$ computes whatever it should compute (the energy, momentum or something else) and finally the particle is created again by $\phi^\dagger$. The integral means that this process is repeated for each point in space and then summed.

Notice also that here is the origin of the awkward term "second quantization". It comes from the fact that the operator ϕ looks like an arbitrary wavefunction but with the coefficients in the expansion replaced by the annihilation operators a_a. Thus it *looks like* the wave function is "quantized again" which of course is not true and is the source of much confusion. Second quantization is just a formalism, within the framework of ordinary quantum theory, to describe systems with many particles and in particular where the types of particles may change.

There is a pleasant surprise incorporated in the second quantized formalism. Since the creation operators $a_a^\dagger$ commute with themselves, any wavefunction is *automatically* symmetric with respect to interchange of these particles as should be the case for bosons. This naturally leads to the question what one should do if one would like like to instead describe fermions since in that case the wavefunctions should be *anti*-symmetric with respect to interchange of particles. The natural thing to try is to use operators which do not commute, but *anti*-commute. That is, operators $b_a, b_a^\dagger$ which satisfy

$$\left\{b_a, b_b^\dagger\right\} = b_a b_b^\dagger + b_b^\dagger b_a = \delta_{ab},$$

$$\{b_a, b_b\} = \left\{b_a^\dagger, b_b^\dagger\right\} = 0. \tag{75}$$

In this case, since the creation operators *anti*-commute, we get an extra minus sign when we interchange particles $b_a^\dagger b_b^\dagger = -b_b^\dagger b_a^\dagger$ giving us the required behavior under interchange of particles. Furthermore we see that if we try to put more than one particle in each state

$$|2\rangle = b^\dagger b^\dagger|0\rangle = \frac{1}{2}\left\{b^\dagger, b^\dagger\right\}|0\rangle = 0! \tag{76}$$

This means that the Pauli principle is automatically incorporated when we use anti-commuting creation/annihilation operators.

6 Dyson's method

Since we have found the time dependent plane wave solutions of the Dirac equation we completely know how time evolution of any state if there are no interaction (the theory without interactions we call the *free* theory). Just Fourier expand the full wavefunction at any given time and then let the individual plane waves evolve in time. The problem comes when we consider an *interacting* theory. Then the Hamiltonian can be written as a sum of two operators $H = H_0 + H_I$ where H_0 is the free Hamiltonian which is responsible for the (almost trivial) time evolution of the free theory (i.e. the plane waves are eigenstates of H_0) and H_I is the interaction Hamiltonian which does not necessarily commute with H_0 or even with itself at different times. This makes the time evolution problem quite involved. However, Dyson has invented a nice little trick which "hides" the (almost trivial) time evolution of the free theory so that we may concentrate on the (slightly more complicated) time evolution given by the interaction Hamiltonian H_I. We want to solve the "Schrödinger" equation

$$\left(H_0 + H_I\right)|\psi\rangle = i\frac{\partial}{\partial t}|\psi\rangle. \tag{77}$$

Now define a new state $|\psi\rangle = e^{-iH_0t}|\chi\rangle$. Inserting this we get

$$\left(H_0 + H_I\right)e^{-iH_0t}|\chi\rangle = i\frac{\partial}{\partial t}\left(e^{-iH_0t}|\chi\rangle\right) = e^{-iH_0t}\left(i\frac{\partial}{\partial t} + iH_0\right)|\chi\rangle. \tag{78}$$

Multiplying from the left with e^{iH_0t} and using that H_0 commutes with itself we get

$$e^{iH_0t}H_I e^{-iH_0t}|\chi\rangle = i\frac{\partial}{\partial t}|\chi\rangle. \tag{79}$$

If we define a time dependent interaction Hamiltonian $H_I(t) = e^{iH_0t}H_I e^{-iH_0t}$ this equation takes a very simple form

$$H_I(t)|\chi\rangle = i\frac{\partial}{\partial t}|\chi\rangle, \tag{80}$$

i.e. it looks just like the Schrödinger equation but with the Hamiltonian $H_I(t)$. The "trivial" time dependence generated by H_0 is taken care of by making the operators in H_I develope in time like in the free theory. Notice that since $|\chi\rangle$ is a solution to the time dependent Schrödinger equation it depends on time as in the so called Schrödinger representation while the operator $H_I(t) = e^{iH_0t}H_I e^{-iH_0t}$ depends on time as an operator in the Heisenberg representation! This is a funny mix of representations known as the interaction representation. Anyway, from our studies of Quantum Mechanics we know how to solve the time dependence of $|\chi\rangle$. The solution is given as

$$|\chi(\mathrm{t})\rangle = U(t,t_0)|\chi(\mathrm{t}_0)\rangle, \tag{81}$$

where we have introduced the time evolution operator

$$U(t,t_0) = T\left[e^{-i\int_{t_0}^{t} H_I(t)dt}\right], \tag{82}$$

where $T\,[]$ represents the time ordering operator.

One may worry that it is really the wave function $|\psi\rangle$ which is the "correct" wave function which one should use to calculate probability amplitudes but it is not difficult to show using what you know about the time evolution operator (exercise) that the expressions

$$\langle\psi_\mathrm{f}|T\left[e^{-i\int(H_0+H_I)dt}\right]|\psi_\mathrm{i}\rangle, \tag{83}$$

and

$$\langle\chi_\mathrm{f}|T\left[e^{-i\int H_I(t)dt}\right]|\chi_\mathrm{i}\rangle, \tag{84}$$

are equal.

7 The second quantized Dirac field

We would now like to write the equivalent of the field operators (67) for the Dirac field. Since we know that electrons are fermions, we know that we should use *anti*-commuting creation/annihilation operators rather than commuting ones. With this in mind we may immediately write down a candidate for the field operators

$$\psi = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \sum_{r=1}^{4} \frac{1}{\sqrt{2\,|E|}} b_{\mathbf{k},r} u^{(r)}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{85}$$

Here $r \in \{1,2,3,4\}$ is an index which runs over the four independent spinor solutions. Since we are using the interaction representation, we now need to make this operator transform in time according to the free theory. That is, $\psi(t) = e^{iH_0 t}\psi e^{-iH_0 t}$. If one then uses the formula $e^A B e^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \dots$ and the fact that $\left[H_0, b^\dagger_{\mathbf{k},r}\right] = E(\mathbf{k}) b^\dagger_{\mathbf{k},r}$[1], one find that $b_{\mathbf{k},r}(t) = b_{\mathbf{k},r} e^{-iEt}$ which tells us that

$$\begin{aligned}\psi(t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \Bigg(&\sum_{r=1}^{2} \frac{1}{\sqrt{2E}} b_{\mathbf{k},r} u^{(r)}(\mathbf{k}) e^{-iE(\mathbf{k})t + i\mathbf{k}\cdot\mathbf{x}} \\ &+ \sum_{r=3}^{4} \frac{1}{\sqrt{2\,|E|}} b_{\mathbf{k},r} u^{(r)}(\mathbf{k}) e^{+i|E(\mathbf{k})|t + i\mathbf{k}\cdot\mathbf{x}} \Bigg). \end{aligned} \tag{86}$$

Notice that we had to separate the positive energy solutions ($r = 1,2$) from the negative energy solutions ($r = 3,4$) since they will have a time dependence of different type! In fact, the time dependence of the negative energy annihilation operators look more like the time dependence of a creation operator. Loosely one can reason as follows; if the operator $\mathcal{O}(t)$ creates something the final state should have bigger energy than the initial state. Thus we have

$$\langle \mathrm{E_f}|\mathcal{O}(t)|\mathrm{E_i}\rangle = \langle \mathrm{E_f}|e^{iH_0 t}\mathcal{O}e^{-iH_0 t}|\mathrm{E_i}\rangle = \langle \mathrm{E_f}|\mathcal{O}e^{i(E_f - E_i)t}|\mathrm{E_i}\rangle, \tag{87}$$

so we see that a creation operator should have the time dependence e^{+iEt}. If, on the other hand, the operator $\mathcal{O}$ would annihilate something, then $E_f < E_i$

[1] This follows from the fact that $H_0|1_{\mathbf{k}}\rangle = E(\mathbf{k})|1_{\mathbf{k}}\rangle$ and that $H_0|0\rangle = 0$. Alternatively one may explicitly evaluate the (second quantized) Hamiltonian which turn out to be $H_0 = \sum_{\mathbf{k},r} E b^\dagger_{\mathbf{k},r} b_{\mathbf{k},r}$

and the time dependence going together with annihilation should thus be of the type e^{-iEt}. This gives us a hint on how to treat the problematic negative energy solutions which we seem to be inherent in any relativistic theory. Namely we define

$$\begin{aligned} d^\dagger_{\mathbf{k},1} &= -b_{-\mathbf{k},4}, \\ d^\dagger_{\mathbf{k},2} &= b_{-\mathbf{k},3}. \end{aligned} \tag{88}$$

The annihilation of a negative energy electron is thus reinterpreted as the *creation* of a new *positive* energy particle, a *positron*. The new particle has exactly the same properties as the electron (mass etc) except that the charge is opposite. It is known as the *anti*-particle of the electron. To go together with this we also redefine the spinors

$$\begin{aligned} v^{(1)}(\mathbf{k}) &= -u^{(4)}(-\mathbf{k}), \\ v^{(2)}(\mathbf{k}) &= u^{(3)}(-\mathbf{k}). \end{aligned} \tag{89}$$

With these new definitions we may rewrite the field operator in the interaction representation as

$$\begin{aligned} \psi(t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \sum_{r=1}^{2} \frac{1}{\sqrt{2E}} \Big(& b_{\mathbf{k},r} u^{(r)}(\mathbf{k}) e^{-iE(\mathbf{k})t + i\mathbf{k}\cdot\mathbf{x}} \\ & + d^\dagger_{\mathbf{k},r} v^{(r)}(\mathbf{k}) e^{+iE(\mathbf{k})t - i\mathbf{k}\cdot\mathbf{x}} \Big) . \end{aligned} \tag{90}$$

According to what we know about second quantization, this operator will annihilate an electron at point $\mathbf{x}$ and create a positron at $\mathbf{x}$.

8 Scattering of electrons in the field of a nucleus (Rutherford/Mott scattering)

We will consider the scattering of an electron in the field of a positively charged heavy particle, typically in the field of a nucleus, so we will take the potential to be:

$$\phi = A_0 = \frac{Ze}{r} \tag{91}$$

The incoming state will of course be a state describing the incoming electron with momentum p and spin r. That is, it will be described by the state

$$|\mathrm{i}\rangle = b^\dagger_{p,r}|0\rangle \tag{92}$$

since the $b^\dagger$ operator creates a state of one electron with the specified momentum and spin. The outgoing, final state is also a state with only one electron, but since the electron has scattered it has a different momentum k and a (possibly) different spin s. This is given by the state

$$|\mathrm{f}\rangle = b^\dagger_{k,s}|0\rangle \tag{93}$$

The probability amplitude for this process (scattering of an electron with momentum p and spin r to an electron with the momentum k and spin s we get by taking the initial state $|\mathrm{i}\rangle$ and evolve it with the time evolution operator and finally taking the overlap with the state $\langle\mathrm{f}|$. The probability amplitude is therefore given by the expression

$$\mathcal{M} = \langle\mathrm{f}|T\ e^{-i\int H_I}|\mathrm{i}\rangle \tag{94}$$

and the probability is of course the absolute square of the probability amplitude.

We know that H_I is small (since e is a small number) so we can evaluate the probability amplitude in perturbation theory

$$\langle\mathrm{f}|T\ e^{-i\int H_I}|\mathrm{i}\rangle = \langle\mathrm{f}|\mathrm{i}\rangle - ie\langle\mathrm{f}|\int d^4x\hat{\bar{\psi}}\not{A}\hat{\psi}|\mathrm{i}\rangle + \ldots \tag{95}$$

Notice here that the time ordering operator T is trivial since all operators are at the same time. Only in the higher order terms is the T operator important. If $k \neq p$ which means that scattering has taken place, the first term is zero.

Inserting the information we have we can compute

$$\begin{aligned}\mathcal{M} \;=\; & -\frac{i}{V}\sum_{q_1,t_1}\sum_{q_2,t_2}\int d^4x\frac{Ze^2}{r}\langle 0|b_{k,s} \\ & \left(\frac{1}{\sqrt{2E_{q_1}}}d_{q_1,t_1}\bar{v}_{q_1,t_1}e^{-iq_1\cdot x} + \frac{1}{\sqrt{2E_{q_1}}}b^\dagger_{q_1,t_1}\bar{u}_{q_1,t_1}e^{iq_1\cdot x}\right) \\ & \gamma^0\left(\frac{1}{\sqrt{2E_{q_2}}}d^\dagger_{q_2,t_2}v_{q_2,t_2}e^{iq_2\cdot x} + \frac{1}{\sqrt{2E_{q_2}}}b_{q_2,t_2}u_{q_2,t_2}e^{-iq_2\cdot x}\right) b^\dagger_{p,r}|0\rangle\end{aligned} \tag{96}$$

There are in principle four different terms but clearly only terms with the same number of creation and annihilation operators will survive. There are two such terms, the first being the one where we select the positron creation/annihilation operators from the parenthesis. The operators squeezed between the vacuum states in that case are

$$\langle 0|b_{k,s}d_{q_1,t_1}d^\dagger_{q_2,t_2}b^\dagger_{p,r}|0\rangle \tag{97}$$

and using the anti-commutation relations we can transform this into

$$\delta_{p,k}\delta_{r,s}\delta_{q_1,q_2}\delta_{t_1,t_2} \tag{98}$$

We see that this term does not give anything unless $p = k$ which means that no scattering is taking place. It would represent a process where the electron is just passing by when a positron is created and annihilated out of the vacuum. This clearly have no effect on the scattering process and we will therefore drop this term.

The second term is the one with only electron creation/annihilation operators which, writing only the operators, gives us

$$\langle 0|b_{k,s}b^\dagger_{q_1,t_1}b_{q_2,t_2}b^\dagger_{p,r}|0\rangle = \delta_{k,q_1}\delta_{s,t_1}\delta_{q_2,p}\delta_{t_2,r} \tag{99}$$

This is the type of term we expect. The interpretation is that the incoming electron gets annihilated and there is a new electron (with new momentum and spin) created, i.e. the electron gets "scattered". We thus get the formula

$$\mathcal{M} = -\frac{i}{V}\sum_{q_1,t_1}\sum_{q_2,t_2}\int d^4x \frac{Ze^2}{r}\frac{e^{i(q_1-q_2)\cdot x}}{\sqrt{4E_{q_1}E_{q_2}}}\left(\bar{u}_{q_1,t_1}\gamma^0 u_{q_2,t_2}\right)\delta_{k,q_1}\delta_{s,t_1}\delta_{q_2,p}\delta_{t_2,r} \tag{100}$$

We can use the Kronecker deltas to get rid of the summations

$$\begin{aligned}\mathcal{M} &= -\frac{i}{V}\int d^4x \frac{Ze^2}{r}\frac{e^{i(k-p)\cdot x}}{\sqrt{4E_pE_k}}\left(\bar{u}_{k,s}\gamma^0 u_{p,r}\right) = \\ &\quad -i\frac{U(\mathbf{k}-\mathbf{p})}{V\sqrt{4E_pE_k}}(\bar{u}_{k,s}\gamma^0 u_{p,r})\int_{-T/2}^{T/2} dt\, e^{i(E_k-E_p)t}\end{aligned} \tag{101}$$

where we have introduced the 3-dimensional Fourier transform of the coulomb potential

$$U(\mathbf{k}) \equiv \int d^3\mathbf{x}\frac{Ze^2}{|\mathbf{x}|}e^{-i\mathbf{k}\cdot\mathbf{x}} = \frac{4\pi Ze^2}{|\mathbf{k}|^2} \tag{102}$$

When we let the interaction time T go to infinity, the last integral in (101) is just (2π times) a delta function of the energy telling us that the energy is conserved in the scattering process.

Now the actual probability, let us call it P, is given by the absolute square of the probability amplitude

$$P = |\mathcal{M}|^2 = \frac{|U|^2}{4V^2 E_p E_k} \left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 |2\pi\delta(E_k - E_p)|^2 \tag{103}$$

Here the last term, the delta function square, may seem a little bit odd, but we can treat it using a trick, namely, we may write it as

$$|2\pi\delta(E_k - E_p)|^2 = \lim_{T\to\infty} 2\pi\delta(E_k - E_p) \int_{-T/2}^{T/2} dt\, e^{i(E_k - E_p)t} \tag{104}$$

Due to the presence of the delta function, the $E_k - E_p$ in the integrand can be replaced by 0. This means that the integrand can be replaced by 1 and thus the integral is equal to T. We therefore get the result

$$P = \lim_{T\to\infty} \frac{|U|^2}{4V^2 E_p E_k} \left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 T 2\pi\delta(E_k - E_p) \tag{105}$$

which gives us an expression for the probability per unit time

$$w \equiv \frac{P}{T} = \frac{|U|^2}{4V^2 E_p E_k} \left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 2\pi\delta(E_k - E_p) \tag{106}$$

Since we cannot experimentally separate scattering into final energy and momenta which are close to each other, we need to sum over all these probabilities to get a total probability for scattering into a state with final energy E_k *or* into a state with energy close to it. This we do by multiplying the probability with the density of states $\rho(E_k)$ and then integrate over energy

$$\begin{aligned} w_{\text{tot}} &= \int dE_k \frac{|U|^2}{4V^2 E_p E_k} \left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 2\pi\delta(E_k - E_p)\rho(E_k) = \\ & \frac{|U|^2}{4V^2 E_p^2} \left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 2\pi\rho(E_p) \end{aligned} \tag{107}$$

This expression now depends only on the energy of the incoming electron E_p which we will hereafter denote by just E. Note however that the spinors

u still depend on the 3-dimensional final momentum $\mathbf{k}$. Because of energy conservation $\mathbf{k} = \mathbf{p}$ but the direction can still be different.

The density of states function $\rho(E)$ we find in the following way. We have assumed that the universe is a (large but finite) box with length, width and hight L. In such a box the allowed 3-momenta are not arbitrary but rather discrete points momentum space. Only momenta with values $k_i = \frac{2\pi}{L} n_i$ $i = 1, 2, 3$, where n_i are integers are allowed. This gives a density of states in momentum space as $\rho(\mathbf{k}) = \frac{L^3}{(2\pi)^3} = \frac{V}{(2\pi)^3}$. Since the energy is a function of the 3-momenta we can write

$$\rho(E)dE = \rho(\mathbf{k})d^3\mathbf{k} = \frac{V}{(2\pi)^3}\mathbf{k}^2 d\left|\mathbf{k}\right| d\Omega_{\mathbf{k}} \tag{108}$$

where we have introduced spherical coordinates in momentum space. From this it follows that

$$\rho(E) = \frac{V}{(2\pi)^3}\mathbf{k}^2 \frac{d\left|\mathbf{k}\right|}{dE} d\Omega_k \tag{109}$$

and since $\frac{dE}{d|\mathbf{k}|} = \frac{|\mathbf{k}|}{E}$ we have

$$\rho(E) = \frac{V}{(2\pi)^3} E\left|\mathbf{k}\right| d\Omega_k \tag{110}$$

We thus have the number

$$w = \frac{|U|^2}{4V^2E^2}\left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 \frac{V}{(2\pi)^2} E\left|\mathbf{k}\right| d\Omega_k \tag{111}$$

representing the probability per unit time that a particle gets scattered into the space angle $d\Omega_k$. More precisely, since the incoming wave-function is extended in all space and is normalized to one, which means that there is only one particle in the whole universe, we have calculated the probability for scattering if we have an incoming flux (= the number of particles per unit time and unit area) of $\frac{v}{V}$ (where $v = \frac{|\mathbf{k}|}{E}$ is the speed of the incoming particle). Since we would like to get a number which is not dependent on the particular incoming flux that we have chosen, we divide the probability w with the flux and get a number called the (differential) *cross section*. This number characterize the physical process and is not dependent on any particular choice of flux used in the experiment. It is given by

$$d\sigma = \frac{|U(\mathbf{k} - \mathbf{p})|^2}{4(2\pi)^2}\left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 d\Omega_k \tag{112}$$

To get the actual number of scattered particles per unit time that we will measure in our detector, we have to multiply this number with the incoming flux we are using in the experiment. If we are interested here in the cross section when the incoming particle has some particular spin and the outgoing particle also has some fixed spin we just insert their corresponding spinors u and $\bar{u}$ in the expression for the cross-section above and we are done.

However, if we assume that the initial state is unpolarized which means that half of the particles have spin up and the other half have spin down, but that the relative phases of the particles are totally random, then the resulting probability (cross section) is given by averaging over the spin of the initial wave-function. In this case that means summing the final result over r and multiplying by $\frac{1}{2}$. If we also do not measure the spin of the outgoing particle we have to sum the final probability (cross section) over the separate probabilities to measure an outgoing particle with spin up and an outgoing particle with spin down. This gives us

$$d\sigma = \frac{|U(\mathbf{k}-\mathbf{p})|^2}{4(2\pi)^2}\frac{1}{2}\sum_r\sum_s\left|\bar{u}_{k,s}\gamma^0 u_{p,r}\right|^2 d\Omega_k \tag{113}$$

The sums over the different spins can be written

$$\sum_r\sum_s\left(\bar{u}_{k,s}\gamma^0 u_{p,r}\bar{u}_{p,r}\gamma^0 u_{k,s}\right) \tag{114}$$

or, writing out the matrix indexes explicitly

$$\left(\sum_s\left(\left(u_{k,s}\right)_a\left(\bar{u}_{k,s}\right)_b\right)\left(\gamma^0\right)_{bc}\sum_r\left(\left(u_{p,r}\right)_c\left(\bar{u}_{p,r}\right)_d\right)\left(\gamma^0\right)_{da}\right) \tag{115}$$

which, defining the matrix $M_{ab}(k) \equiv \sum_s \left(u_{k,s}\right)_a\left(\bar{u}_{k,s}\right)_b$, can be written

$$\mathrm{Tr}\left(M(k)\gamma^0 M(p)\gamma^0\right) \tag{116}$$

Using the explicit representation of the spinors one can find that

$$M(k) = \not{k} + m \tag{117}$$

so we have

$$d\sigma = \frac{|U(\mathbf{k}-\mathbf{p})|^2}{8(2\pi)^2}\mathrm{Tr}\left((\not{k}+m)\gamma^0(\not{p}+m)\gamma^0\right) d\Omega_k \tag{118}$$

Using the gamma-matrix anti-commutation relations we can compute

$$\mathrm{Tr}\left((\not{k}+m)\gamma^0(\not{p}+m)\gamma^0\right) = 4\left(m^2 + E_k E_p + \mathbf{k}\cdot\mathbf{p}\right) \tag{119}$$

Finally using that $E_k = E_p$ and $|\mathbf{k}| = |\mathbf{p}|$ we can choose a coordinate system so that $\mathbf{p}$ is along the z-axis and $\mathbf{k}$ is pointing in the (θ, ϕ) direction. Inserting this we get the final formula for the relativistically corrected Rutherford formula also called the Mott cross section.

$$d\sigma = \frac{Z^2 e^4}{4\sin^4\left(\frac{\theta}{2}\right)} \frac{E^2}{|\mathbf{k}|^4}\left(1 - v^2 \sin^2\left(\frac{\theta}{2}\right)\right) d\Omega \tag{120}$$

9 Pair creation

What is the probability that an electron/positron pair is created in the potential

$$A_\mu = \left(0, 0, 0, \sqrt{4\pi}a\cos(\omega t)\right) \tag{121}$$

This 4-potential represents an electric field directed in the 3 direction and oscillating with frequency ω. In this case the initial and final states are of course given by

$$\begin{aligned} |\mathrm{i}\rangle &= |0\rangle \\ |\mathrm{f}\rangle &= b^\dagger_{k_1,r_1} d^\dagger_{k_2,r_2}|0\rangle \end{aligned} \tag{122}$$

representing the fact that initially we do not have any particles at all but we will end up with both an electron with momentum k_1 and a positron with momentum k_2. As usual the probability amplitude is given by

$$\mathcal{M} = \langle \mathrm{f}|T\ e^{-i\int H_I}|\mathrm{i}\rangle \tag{123}$$

which, to lowest order in the expansion parameter e can be written as

$$\mathcal{M} = -ie\langle \mathrm{f}|\int d^4x \hat{\bar{\psi}} \not{A} \hat{\psi}|0\rangle = -ie\int d^4x A_3 \langle 0|b_{k_1,r_1} d_{k_2,r_2} \hat{\bar{\psi}}\gamma^3\hat{\psi}|0\rangle \tag{124}$$

It is quite clear that the only piece that will survive is the piece containing the operator $b^\dagger$ from $\hat{\bar{\psi}}$ and the operator $d^\dagger$ from $\hat{\psi}$. What remains is

$$-\frac{ie}{V}\sum_{q_1,s_1}\sum_{q_2,s_2}\int d^4x A_3 \frac{e^{i(q_1+q_2)\cdot x}}{\sqrt{4E_{q_1}E_{q_2}}}\left(\bar{u}_{q_1,s_1}\gamma^3 v_{q_2,s_2}\right)\langle 0|b_{k_1,r_1} d_{k_2,r_2} b^\dagger_{q_1,s_1} d^\dagger_{q_2,s_2}|0\rangle \tag{125}$$

which, using the anti-commutation relations simplifies to

$$\frac{ie}{V}\int d^4x A_3 \frac{e^{i(k_1+k_2)\cdot x}}{\sqrt{4E_{k_1}E_{k_2}}}\left(\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right) \tag{126}$$

Inserting the expression for the potential we write

$$\frac{iea\sqrt{4\pi}}{V\sqrt{4E_{k_1}E_{k_2}}}\left(\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right)\int d^4x\cos(\omega t)e^{i(k_1+k_2)\cdot x} \tag{127}$$

The integral can be performed by rewriting $\cos(\omega t)$ in terms of exponentials as

$$\int d^4x\cos(\omega t)e^{i(k_1+k_2)\cdot x} = \int d^3\mathbf{x}e^{-i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{x}}\int dte^{i(E_{k_1}+E_{k_2})t}\frac{e^{i\omega t}+e^{-i\omega t}}{2} = \frac{(2\pi)^4}{2}\delta^3\left(\mathbf{k}_1+\mathbf{k}_2\right)\left(\delta\left(E_{k_1}+E_{k_2}+\omega\right)+\delta\left(E_{k_1}+E_{k_2}-\omega\right)\right) \tag{128}$$

The term containing $\delta\left(E_{k_1}+E_{k_2}+\omega\right)$ will clearly not give any contribution since E_{k_1}, E_{k_2} and ω are all positive. Therefore we have for the probability amplitude

$$\mathcal{M} = \frac{i(2\pi)^4 ea\sqrt{4\pi}}{2V\sqrt{4E_{k_1}E_{k_2}}}\left(\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right)\delta^3\left(\mathbf{k}_1+\mathbf{k}_2\right)\delta\left(E_{k_1}+E_{k_2}-\omega\right) \tag{129}$$

Again we see that the delta functions express energy conservation $\omega = E_{k_1} + E_{k_2}$ and momentum conservation $\mathbf{k}_1+\mathbf{k}_2 = 0$. Namely, the frequency of the electric field has to represent an energy which precisely matches the energy of the created electron/positron pair. Also, the electron/positron has to come out back-to-back so that momentum is conserved. This also means that $E_{k_1} = E_{k_2} \equiv E$.

The probability is the absolute square of the probability amplitude

$$P = \frac{\pi e^2 a^2}{4V^2E^2}\left|\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right|^2\left|(2\pi)^3\delta^3\left(\mathbf{k}_1+\mathbf{k}_2\right)(2\pi)\delta\left(2E-\omega\right)\right|^2 \tag{130}$$

By a similar trick as in the last section we evaluate the square of the delta functions to be

$$\left|(2\pi)^3\delta^3\left(\mathbf{k}_1+\mathbf{k}_2\right)(2\pi)\delta\left(2E-\omega\right)\right|^2 = VT(2\pi)^4\delta^3\left(\mathbf{k}_1+\mathbf{k}_2\right)\delta\left(2E-\omega\right) \tag{131}$$

giving us

$$P = T\frac{\pi e^2 a^2 (2\pi)^4}{4VE^2} \left|\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right|^2 \delta^3\left(\mathbf{k}_1 + \mathbf{k}_2\right)\delta\left(2E - \omega\right) \tag{132}$$

As in the previous example, we cannot separate final states which are too close in phase space. Therefore we have to sum over these probabilities to get a total "effective" probability. As in the previous example, this means including a factor $\frac{V}{(2\pi)^3}d^3\mathbf{k}$ for each final particle giving us

$$P = T\frac{\pi e^2 a^2 (2\pi)^4}{4VE^2} \left|\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right|^2 \delta^3\left(\mathbf{k}_1 + \mathbf{k}_2\right)\delta\left(2E - \omega\right)\frac{V}{(2\pi)^3}d^3\mathbf{k}_1\frac{V}{(2\pi)^3}d^3\mathbf{k}_2 \tag{133}$$

One of the integrals is easily performed using the first delta function giving us that $\mathbf{k}_1 = -\mathbf{k}_2 \equiv \mathbf{k}$ and the second integral we can perform, as in the previous case after rewriting $d^3\mathbf{k} = |\mathbf{k}|^2 d|\mathbf{k}|\, d\Omega_k = |\mathbf{k}|^2 \frac{d|\mathbf{k}|}{dE} dE d\Omega_k = |\mathbf{k}|\, E dE d\Omega_k$. The result is (noticing that $\delta(2E-\omega) = \frac{1}{2}\delta(E - \frac{\omega}{2})$)

$$P = VT\frac{e^2 a^2 |\mathbf{k}|}{16\pi\omega}\left|\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right|^2 d\Omega_k \tag{134}$$

Notice that the probability is proportional to the volume (of the region with the electric field) and the time we let the field act, in accordance to physical expectations. This makes it more useful to speak about the probability per unit volume and unit time, $\frac{P}{VT}$.

If we are not observing the spins of the final particles we have to sum over the different probabilities of observing the different possible spins. Then we will get the formula

$$\frac{P}{VT} = \frac{e^2 a^2 |\mathbf{k}|}{16\pi\omega} d\Omega_k \sum_{r_1,r_2}\left|\bar{u}_{k_1,r_1}\gamma^3 v_{k_2,r_2}\right|^2 \tag{135}$$

Using the trick of the last section, this can be rewritten as

$$\frac{P}{VT} = \frac{e^2 a^2 |\mathbf{k}|}{16\pi\omega} d\Omega_k \mathrm{Tr}\left((\not{k}_1 + m)\gamma^3(\not{k}_2 - m)\gamma^3\right) \tag{136}$$

or, using the anti commutation relations of the gamma matrices and remembering that $\mathbf{k}_1 = (E, \mathbf{k})$, $\mathbf{k}_2 = (E, -\mathbf{k})$

$$\begin{aligned}\frac{P}{VT} &= \frac{e^2 a^2 |\mathbf{k}|}{16\pi\omega} d\Omega_k 2\omega^2\left(1 - v^2\cos^2(\theta)\right) = \\ &\frac{e^2 a^2 \omega^2}{16\pi}\sqrt{1 - \frac{4m^2}{\omega^2}}\left(1 - v^2\cos^2(\theta)\right) d\Omega_k\end{aligned} \tag{137}$$

which gives the probability of an electron/positron pair with a momentum with angle θ towards the electric field. If we are interested in the total probability, irrespective of the angle, we have to integrate over $d\Omega$ to get

$$\frac{P}{VT} = \frac{e^2 a^2 \omega^2}{6} \sqrt{1 - \frac{4m^2}{\omega^2}} \left(1 + \frac{2m^2}{\omega^2} \right) \tag{138}$$

Notice that there is a "threshold" in the energy. The probability is zero for $\omega \leq 2m \leq 2E$, i.e. the energy of the photons in the field must be larger than the mass of the electron/positron pair to be able to create it. In contrast the probability is non-zero for arbitrary small amplitude a of the field.

10 The quantized electro-magnetic field

In the two previous examples the elecro-magnetic field was treated classically as an external field. In order to incorporate photons into the theory we need to quantize also the electro-magnetic field. We will do this in a relativistically covariant fashion so let us start by recapitulating some notation. Remember that we may use a scalar potential ϕ and a vector potential $\mathbf{A}$ to describe the electric and magnetic fields

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \partial_t \mathbf{A}, \\ \mathbf{B} &= \nabla \times \mathbf{A}. \end{aligned} \tag{139}$$

Introducing the four vector $A^\mu = (\phi, \mathbf{A})$ we may write the electric and the magnetic field in a compact form

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{140}$$

Using $F^{\mu\nu}$, Maxwell's equations can also be written covariantly as

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= 0, \\ \partial^\mu F^{\nu\sigma} + \partial^\sigma F^{\mu\nu} + \partial^\nu F^{\sigma\mu} &= 0. \end{aligned} \tag{141}$$

Here we have used Einstein's summation convention (sum over repeated indexes). It is interesting to observe that A^μ is not uniquely specified by the electric and magnetic field. Namely, if we define a new vector potential by

$$A^\mu_{\text{new}} = A^\mu_{\text{old}} + \partial^\mu \chi, \tag{142}$$

for any function $\chi(x)$, the field strength $F^{\mu\nu}$, and thus the electric and magnetic fields, remain unchanged. This we can use to simplify the form of Maxwell's equations. If we choose χ so that $\partial_\mu A^\mu_{\rm new} = 0$, i.e. so that $\partial_\mu \partial^\mu \chi + \partial_\mu A^\mu_{\rm old} = 0$, then we have

$$\partial_\mu F^{\mu\nu}_{\rm new} = \partial_\mu \left(\partial^\mu A^\nu_{\rm new} - \partial^\nu A^\mu_{\rm new} \right) = \partial_\mu \partial^\mu A^\nu_{\rm new} = 0, \tag{143}$$

that is, each component of the vector potential has to satisfy the (massless) Klein-Gordon equation (which we have already solved!). Thus, the vector potential we will use will have to satisfy two equations

$$\begin{aligned} \partial_\mu A^\mu &= 0, \\ \partial_\mu \partial^\mu A^\nu &= 0. \end{aligned} \tag{144}$$

However, this still does not completely specify A^μ. We may still shift it as $A^\mu_{\rm fin} = A^\mu_{\rm new} + \partial^\mu \Lambda$ with a Λ satisfying $\partial_\mu \partial^\mu \Lambda = 0$ since this leaves the two equations (144) invariant. This additional invariance can be used to choose $A^0_{\rm fin} = 0$. We have thus seen that we can always choose a vector potential which satisfies the following three equations

$$\begin{aligned} A^0 &= 0, \\ \partial_\mu A^\mu &= 0, \\ \partial_\mu \partial^\mu A^\nu &= 0. \end{aligned} \tag{145}$$

This choice of form of the vector potential (or choice of gauge as the jargon goes) is known as the Coulomb gauge.

Using this information we may now immediately write down the second quantized electromagnetic field

$$A_\mu = \frac{1}{\sqrt{V}} \sum_{\mathbf{k},\alpha} \frac{1}{\sqrt{2\omega}} \left(a_{k,\alpha} \epsilon^{(\alpha)}_\mu e^{-ikx} + a^\dagger_{k,\alpha} \epsilon^{\star(\alpha)}_\mu e^{ikx} \right). \tag{146}$$

From the third equation in (145) we find that $k_\mu k^\mu = \omega^2 - \mathbf{k}^2 = 0$. The second equation tells us that $k \cdot \epsilon^{(\alpha)} = 0$ while the first equation tells us that $\epsilon^{(\alpha)}_0 = 0$. We thus find that $\epsilon^{(\alpha)}_\mu$ is a four vector with zero time component and orthogonal to the four momentum. Thus, out of the four orthonormal four vectors

$$\epsilon^{(0)}_\mu \;=\; (1, \mathbf{0}),$$

$$
\begin{aligned}
\epsilon_\mu^{(1)} &= (0, \bar{\epsilon}^{(1)}), \\
\epsilon_\mu^{(2)} &= (0, \bar{\epsilon}^{(2)}), \\
\epsilon_\mu^{(3)} &= (0, \frac{\mathbf{k}}{|\mathbf{k}|}),
\end{aligned} \tag{147}
$$

only $\epsilon_\mu^{(1)}$ and $\epsilon_\mu^{(2)}$ are admissible. Thus we see that the sum over α in (146) is restricted to $\alpha = 1, 2$ in order for A_μ to satisfy all three equations in (145). The $\epsilon_\mu^{(\alpha)}$ vectors are known as the polarization vectors of the photon. As usual, when we second quantize, the $a_{\mathbf{k},\alpha}, a_{\mathbf{k},\alpha}^\dagger$ become annihilation and creation operators which annihilate/create a photon with momentum k and polarization $\epsilon_\mu^{(\alpha)}$.

11 The electron propagator

If we are interested in the next to lowest order corrections to the scattering in an external potential, we have to study the term

$$
-\frac{e^2}{2}\langle \mathrm{f}| \int d^4x_1 \int d^4x_2 T \left[(\bar{\psi}\not{A}\psi)(x_1)(\bar{\psi}\not{A}\psi)(x_2) \right] |\mathrm{i}\rangle \tag{148}
$$

Notice here that the T operator is non-trivial and important since the two $\bar{\psi}\not{A}\psi$ factors changes place if $t_2 > t_1$.

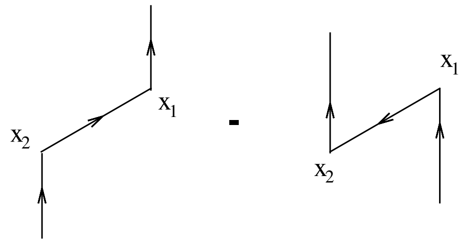
Let us first assume that $t_1 > t_2$. Then the ordering given above is the correct one and we can use the term as it stands. For each ψ factor there are two different choices for the operator, one associated to the electron and one associated with the positron. Since there are four ψ operators, we have in principle 16 different terms. However, out of these 16 terms, only 2 are non-trivial. All the others are either zero or they represent "non-connected" terms in the sense discussed before. For instance, there is one term which represents an electron getting scattered at x_1 while at x_2 a positron is created and annihilated but there is no contact between these two points. The first non-trivial term is schematically

$$
\langle 0| b_f b_1^\dagger b_1 b_2^\dagger b_2 b_i^\dagger |0\rangle = \delta_{f1}\delta_{12}\delta_{2i} \tag{149}
$$

representing the incoming electron being scattered first at x_2 and then, later at x_1. The second non-trivial term is

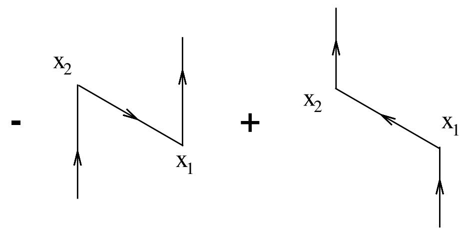
$$
\langle 0| b_f d_1 b_1 b_2^\dagger d_2^\dagger b_i^\dagger |0\rangle = -\delta_{12}\delta_{f2}\delta_{1i} + \delta_{12}\delta_{12}\delta_{fi} \tag{150}
$$

the second term here is again "non-connected", representing a process with no scattering because of the δ_{fi} but the first term is interesting and slightly hard to interpret. Since $t_1 > t_2$ we have to interpret it as the incoming electron flies past the point x_2 where there is an electron/positron pair created and later annihilates with the positron just created at x_2. The electron created at x_1 is in fact the final outgoing electron. Graphically we have



(1) (2)

Similarly, when $t_2 > t_1$ we have graphically



(3) (4)

Using this knowledge we write diagram (1) and (3) (where the initial electron is annihilated in x_2) as

$$\langle 0|b_f\bar{\psi}_a(x_1)|0\rangle\langle 0|T\left[\psi_b(x_1)\bar{\psi}_c(x_2)\right]|0\rangle\langle 0|\psi_d(x_2)b_i^\dagger|0\rangle \tag{151}$$

and the diagrams (2) and (4) (where the initial electron is annihilated in x_1) as

$$\langle 0|b_f\bar{\psi}_c(x_2)|0\rangle\langle 0|T\left[\psi_d(x_2)\bar{\psi}_a(x_1)\right]|0\rangle\langle 0|\psi_b(x_1)b_i^\dagger|0\rangle \tag{152}$$

and the total amplitude is of course the sum of these two terms. We see that in this expression the object $\langle 0|T\left[\psi_a(x_1)\bar{\psi}_b(x_2)\right]|0\rangle$ plays an important role. It represents the particle going between the points x_1 and x_2 and if $t_1 > t_2$ it is an electron but if $t_2 > t_1$ it is a positron. This object is therefore called *the electron/positron propagator* and we will now proceed to calculate it.

Because of the time ordering operator T we have to consider two cases. Assume to begin with that $t_1 > t_2$. Then we pick out the electron creation/annihilation operators and the propagator can be written as

$$\sum_{k_1,s_1}\sum_{k_2,s_2}\frac{e^{-ik_1\cdot x_1+ik_2\cdot x_2}}{2V\sqrt{E_1E_2}}(u_1)_a(\bar{u}_2)_b\langle 0|b_1b_2^\dagger|0\rangle \tag{153}$$

Since $\langle 0|b_1b_2^\dagger|0\rangle = \delta_{k_1,k_2}\delta_{s_1,s_2}$ we can evaluate one of the sums "for free"

$$\sum_{k_1,s_1}\frac{e^{ik_1\cdot(x_2-x_1)}}{2VE_1}(u_1)_a(\bar{u}_1)_b \tag{154}$$

Using the results of the previous section ($\sum_s u_a^{(s)}\bar{u}_b^{(s)} = (\not{k}+m)_{ab}$) we can calculate the sum over the spin in the expression for the propagator. The result is

$$\sum_{k_1}\frac{e^{ik_1\cdot(x_2-x_1)}}{2VE_1}(\not{k}_1+m)_{ab} \tag{155}$$

For convenience we will here change the summation over momenta $\sum_k$ into an integral. This we can do since the volume of the universe is large (so that the distribution of states in momentum space is almost continuous) which means that $\sum_k = \frac{V}{(2\pi)^3}\int d^3\mathbf{k}$. This gives us

$$\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{iE(t_2-t_1)}e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)}\frac{(\not{k}+m)_{ab}}{2E} \tag{156}$$

(where we have dropped the index 1 for convenience). Notice here that E is a function of $\mathbf{k}$.

The same analysis in the case where $t_2 > t_1$ gives

$$-\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{iE(t_1-t_2)}e^{-i\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_2)}\frac{(\not{k}-m)_{ab}}{2E} \tag{157}$$

where the minus sign comes from the fact that the T operator has reordered two *fermionic* operators.

We can rewrite the result in a more covariant form by using the integral

$$-\frac{e^{-iE|t|}}{2E} = \lim_{\varepsilon\to 0}\frac{1}{2\pi i}\int_{-\infty}^{\infty}dk_0\frac{e^{-ik_0t}}{k_0^2-E^2+i\varepsilon} \tag{158}$$

The first part of the propagator then becomes

$$
\begin{aligned}
\int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{iE(t_2-t_1)} e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)} \frac{(\not{k}+m)_{ab}}{2E} &= \\
(-i\not{\partial}_2+m)_{ab} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{e^{iE(t_2-t_1)}}{2E} e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)} &= \\
(-i\not{\partial}_2+m)_{ab} \int \frac{d^3\mathbf{k}}{(2\pi)^3} i \int \frac{dk_0}{2\pi} \frac{e^{ik_0(t_2-t_1)}}{k_0^2-E^2+i\varepsilon} e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)} & \qquad (159)
\end{aligned}
$$

Since $E = \mathbf{k}^2 + m^2$ the denominator of the integrand can be written $k_0^2 - E^2 + i\varepsilon = k_0^2 - \mathbf{k}^2 - m^2 + i\varepsilon = k^2 - m^2 + i\varepsilon$ giving us

$$
i\,(-i\not{\partial}_2+m)_{ab} \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot(x_2-x_1)}}{k^2-m^2+i\varepsilon} \qquad (160)
$$

and pushing the derivative operator back in, we get

$$
i \int \frac{d^4k}{(2\pi)^4} \frac{(\not{k}+m)_{ab}}{k^2-m^2+i\varepsilon} e^{ik\cdot(x_2-x_1)} \qquad (161)
$$

The second part of the propagator (the one with $t_2 > t_1$) can similarly be rewritten

$$
i \int \frac{d^4k}{(2\pi)^4} \frac{(-\not{k}+m)_{ab}}{k^2-m^2+i\varepsilon} e^{ik\cdot(x_1-x_2)} \qquad (162)
$$

and, changing the integration variable from k to $-k$, we get

$$
i \int \frac{d^4k}{(2\pi)^4} \frac{(\not{k}+m)_{ab}}{k^2-m^2+i\varepsilon} e^{ik\cdot(x_2-x_1)} \qquad (163)
$$

which is exactly the same expression as for the part of the propagator with $t_1 > t_2$. We thus have a unique expression for the propagator

$$
G(x_1-x_2) \equiv i \int \frac{d^4k}{(2\pi)^4} \frac{(\not{k}+m)_{ab}}{k^2-m^2+i\varepsilon} e^{-ik\cdot(x_1-x_2)} \qquad (164)
$$

independent of whether t_1 or t_2 comes first.

Notice that if we act with the Dirac operator $i\not{\partial} - m$ on the propagator we get

$$
\begin{aligned}
(i\not{\partial}-m)G(x) = i \int \frac{d^4k}{(2\pi)^4} \frac{(\not{k}-m)(\not{k}+m)}{k^2-m^2+i\varepsilon} e^{-ik\cdot x} &= \\
i \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot x} = i\delta^4(x) & \qquad (165)
\end{aligned}
$$

so that $G(x)$ is the kernel of the Dirac operator in accordance with the usual interpretation of the propagator.

12 Compton scattering

Compton scattering is scattering of a photon on an electron. In the initial state we therefore have a photon and an electron

$$|\mathrm{i}\rangle = b^\dagger_{p,s} a^\dagger_{k,\alpha}|0\rangle \tag{166}$$

Since the photon is physical the index α takes only the values $1, 2$ corresponding to the two physical polarizations of the photon. The final state also contains a photon and an electron but with different spins and momenta

$$|\mathrm{f}\rangle = b^\dagger_{p',s'} a^\dagger_{k',\alpha'}|0\rangle \tag{167}$$

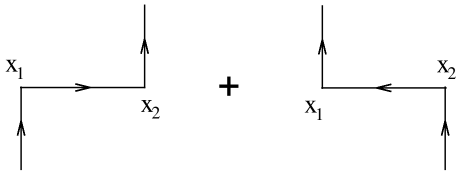
As usual, the probability amplitude is given by

$$\mathcal{M} = \langle \mathrm{f}|T\, e^{-i\int H_I}|\mathrm{i}\rangle = \langle \mathrm{f}|\mathrm{i}\rangle - i\langle \mathrm{f}|\int H_I|\mathrm{i}\rangle - \frac{1}{2}\langle \mathrm{f}|T\int H_I \int H_I|\mathrm{i}\rangle + \ldots \tag{168}$$

The first term is non-zero only for the case when no scattering is taking place ($p = p'$, $k = k'$). The second term is zero because it always involves exactly three photon creation/annihilation operators so the lowest non-trivial term is the third one. Separating the piece that has to do with electrons/positrons and the piece that has to do with photons we can write it as

$$\begin{aligned} -\frac{e^2}{2}\int d^4x_1 d^4x_2 \langle 0|b_{p',s'}T\left[(\bar\psi_a\psi_b)(x_1)(\bar\psi_c\psi_d)(x_2)\right] b^\dagger_{p,s}|0\rangle \\ (\gamma^\mu)_{ab}\,(\gamma^\nu)_{cd}\,\langle 0|a_{k',\alpha'}T\left[A_\mu(x_1)A_\nu(x_2)\right] a^\dagger_{k,\alpha}|0\rangle \end{aligned} \tag{169}$$

The piece associated with the electrons/positrons we have already calculated in the previous section. It is given by two terms corresponding to the cases where the electron first goes to the point x_2, interacts, then goes to the point x_1 where it is scattered to the final electron state and, oppositely when it goes first to x_1 and then continues to x_2. Graphically this can be represented as



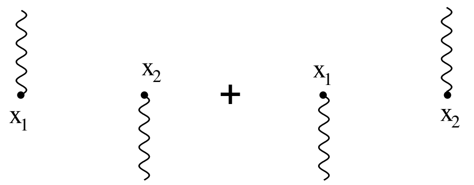
For the second process the expression is

$$\frac{(\bar{u}_{p',s'})_a}{\sqrt{2VE'}} e^{ip'\cdot x_1} \; i \int \frac{d^4 q}{(2\pi)^4} \frac{(\not{q} + m)_{bc}}{q^2 - m^2 + i\varepsilon} e^{-iq\cdot(x_1 - x_2)} \; \frac{(u_{p,s})_d}{\sqrt{2VE}} e^{-ip\cdot x_2} \tag{170}$$

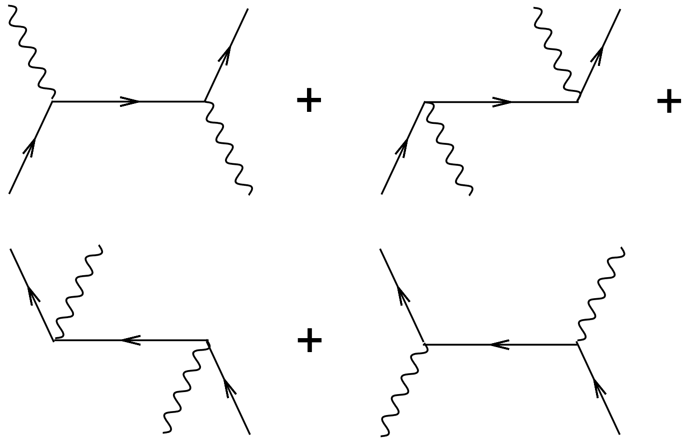
Here we recognize the middle part as the electron/positron propagator corresponding to the piece where the electron propagates from the point x_2 to the point x_1. The expression for the first term is similar. Evaluating the photon piece we get a sum of two terms. The first comes from choosing the creation operator in $A_\mu(x_1)$ and the annihilation operator in $A_\nu(x_2)$, representing the case where the incoming photon is annihilated in x_2 and the outgoing photon is created in x_1, and the second term (where we also have to ignore a "non-connected" piece) comes from choosing the annihilation operator in $A_\mu(x_1)$ and the creation operator in $A_\nu(x_2)$, representing the case where the incoming photon is annihilated in x_1 and the outgoing photon is created in x_2. Explicitly we have

$$4\pi \frac{(\varepsilon^\star_{k',\alpha'})_\mu}{\sqrt{2V\omega'}} e^{ik'\cdot x_1} \frac{(\varepsilon_{k,\alpha})_\nu}{\sqrt{2V\omega}} e^{-ik\cdot x_2} + 4\pi \frac{(\varepsilon^\star_{k',\alpha'})_\nu}{\sqrt{2V\omega'}} e^{ik'\cdot x_2} \frac{(\varepsilon_{k,\alpha})_\mu}{\sqrt{2V\omega}} e^{-ik\cdot x_1} \tag{171}$$

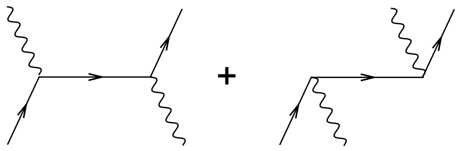
Graphically we can write this as



Putting these two terms together we graphically get



Since we are integrating over x_1 and x_2 in the final expression these variables are really "dummy" variables, meaning that we can anywhere rename them as we wish. In particular we can interchange them $x_1 \leftrightarrow x_2$. From the pictures we see that two of the pictures change into the other two under this relabeling so we have really to calculate only two terms, graphically they look like this



The second term we can write as

$$\frac{\pi}{V^2\sqrt{E'E\omega'\omega}}\int d^4x_1\int d^4x_2 i\int\frac{d^4q}{(2\pi)^4}\frac{\bar{u}_{p',s'}\not{\varepsilon}_{k',\alpha'}(\not{q}+m)\not{\varepsilon}_{k,\alpha}u_{p,s}}{q^2-m^2+i\varepsilon}$$
$$e^{ip'\cdot x_2}e^{ik'\cdot x_2}e^{-ip\cdot x_1}e^{-ik\cdot x_1}e^{-iq\cdot(x_2-x_1)} \qquad (172)$$

and doing the integrals over x_1 and x_2 we get

$$\frac{\pi}{V^2\sqrt{E'E\omega'\omega}}i\int\frac{d^4q}{(2\pi)^4}\frac{\bar{u}_{p',s'}\not{\varepsilon}_{k',\alpha'}(\not{q}+m)\not{\varepsilon}_{k,\alpha}u_{p,s}}{q^2-m^2+i\varepsilon}$$
$$(2\pi)^4\delta^4(p'+k'-q)(2\pi)^4\delta^4(q-p-k) \qquad (173)$$

Notice that the delta functions express momentum conservation at each of the vertexes. The expression for the first diagram is the same except that one have to switch positions for the polarization vectors ε and switch the place of k and k' in the delta functions. We can get rid of one of the delta functions by performing the q integral which gives us

$$-\frac{i\pi e^2(2\pi)^4\delta^4(p'+k'-p-k)}{V^2\sqrt{EE'\omega\omega'}}\left[\bar{u}_{p',s'}\not{\varepsilon}_{k',\alpha'}\frac{\not{p}+\not{k}+m}{(p+k)^2-m^2+i\varepsilon}\not{\varepsilon}_{k,\alpha}u_{p,s}+\right.$$
$$\left.\bar{u}_{p',s'}\not{\varepsilon}_{k,\alpha}\frac{\not{p}-\not{k}'+m}{(p-k')^2-m^2+i\varepsilon}\not{\varepsilon}_{k',\alpha'}u_{p,s}\right] \qquad (174)$$

Here we notice that the delta function which is left just expresses the momentum conservation of the whole process. To simplify this expression further we can use the Dirac equation on the spinors

$$\not{p}u_p = mu_p \qquad (175)$$

which, using the anti-commutation relations of the gamma matrices, leads to

$$\not{p}\not{\varepsilon}u_p = -\not{\varepsilon}mu_p + 2p\cdot\varepsilon u_p \tag{176}$$

and choosing a coordinate system where the initial electron is at rest $p = (m,0,0,0)$ so that $p\cdot\varepsilon = p\cdot\varepsilon' = 0$ we get

$$-\frac{i\pi e^2(2\pi)^4\delta^4(p'+k'-p-k)}{V^2\sqrt{EE'\omega\omega'}}\left[\frac{\bar{u}'\not{\varepsilon}'\not{k}\not{\varepsilon}u}{2p\cdot k}+\frac{\bar{u}'\not{\varepsilon}\not{k}'\not{\varepsilon}'u}{2p\cdot k'}\right] \tag{177}$$

Now we calculate the probability density by taking the absolute square of the amplitude. At the same time we say that we are not interested in the polarization of the final electron, and the initial electron comes in a mixed state so that we have to include a sum $\frac{1}{2}\sum_{s,s'}$. We then have the probability

$$\begin{aligned} P &= \left|\frac{\pi e^2(2\pi)^4\delta^4(p'+k'-p-k)}{2V^2\sqrt{EE'\omega\omega'}}\right|^2 \\ & \frac{1}{2}\sum_{s,s'}\left(\frac{\bar{u}'\not{\varepsilon}'\not{k}\not{\varepsilon}u}{p\cdot k}+\frac{\bar{u}'\not{\varepsilon}\not{k}'\not{\varepsilon}'u}{p\cdot k'}\right)\left(\frac{\bar{u}\not{\varepsilon}\not{k}\not{\varepsilon}'u'}{p\cdot k}+\frac{\bar{u}\not{\varepsilon}'\not{k}'\not{\varepsilon}u'}{p\cdot k'}\right) \end{aligned} \tag{178}$$

where we have used that $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$. Performing the spin sums and noticing that the first term in each parenthesis is equal to the second if we make the exchange $\varepsilon \leftrightarrow \varepsilon'$ and $k \leftrightarrow -k'$ we have

$$\begin{aligned} P &= \left|\frac{\pi e^2(2\pi)^4\delta^4(p'+k'-p-k)}{2V^2\sqrt{2EE'\omega\omega'}}\right|^2 \\ & \mathrm{Tr}\left(\frac{\not{\varepsilon}'\not{k}\not{\varepsilon}(\not{p}+m)\not{\varepsilon}\not{k}\not{\varepsilon}'(\not{p}'+m)}{(p\cdot k)^2}+\right. \\ & \left.\frac{\not{\varepsilon}'\not{k}\not{\varepsilon}(\not{p}+m)\not{\varepsilon}'\not{k}'\not{\varepsilon}(\not{p}'+m)}{(p\cdot k)(p\cdot k')}+\left\{\begin{array}{ccc}\varepsilon & \leftrightarrow & \varepsilon' \\ k & \leftrightarrow & -k'\end{array}\right\}\right) \end{aligned} \tag{179}$$

The reason for the minus sign in the exchange of the photon momenta is that we do not want p or p' to change, but since $p' = p + k - k'$ we have to interchange k and k' with an extra minus sign.

Let us perform the trace over the first term explicitly. The fact that the trace over any odd number of gamma matrices is zero allows us to write it as

$$\mathrm{Tr}\,(\not{\varepsilon}'\not{k}\not{\varepsilon}\not{p}\not{\varepsilon}\not{k}\not{\varepsilon}'\not{p}') + m^2\mathrm{Tr}\,(\not{\varepsilon}'\not{k}\not{\varepsilon}\not{\varepsilon}\not{k}\not{\varepsilon}') \tag{180}$$

Using the gamma matrix algebra we know that $\not a \not b = -\not b \not a + 2a \cdot b$ which also implies that $\not a \not a = a \cdot a$, and using the cyclicity of the trace, we can show that the second term is

$$m^2 \mathrm{Tr}\,(\not k \not \varepsilon \not \varepsilon \not k \not \varepsilon' \not \varepsilon') = m^2 (\varepsilon \cdot \varepsilon)(\varepsilon' \cdot \varepsilon') \mathrm{Tr}\,(\not k \not k) = 0 \tag{181}$$

Thus it only remains to calculate the first term. Anti-commuting the $\not p \not k$ in the middle and using that $\not k \not k = k \cdot k = 0$ we have

$$\mathrm{Tr}\,(\not \varepsilon' \not k \not \varepsilon \not p \not \varepsilon \not k \not \varepsilon' \not p') = 2(p \cdot k) \mathrm{Tr}\,(\not \varepsilon' \not k \not \varepsilon' \not p') \tag{182}$$

Then we can anti-commute $\not \varepsilon' \not k$ and use that $\varepsilon' \cdot \varepsilon' = -1$ to get

$$2(k \cdot p) \mathrm{Tr}\,((-\not k \not \varepsilon' + 2(\varepsilon' \cdot k)) \not \varepsilon' \not p') = 8(k \cdot p)\,(2(k \cdot \varepsilon')(\varepsilon' \cdot p') + (k \cdot p)) \tag{183}$$

where we in the last step used that $\mathrm{Tr}\,(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu}$.

Similarly the second term in the full trace can be calculated to be

$$-8(k' \cdot p)(k \cdot p) + 16(\varepsilon' \cdot \varepsilon)^2 (k' \cdot p)(k \cdot p) - 8(k \cdot \varepsilon')^2 (k' \cdot p) + 8(k' \cdot \varepsilon)^2 (k \cdot p) \tag{184}$$

and the two last terms are given by the first two by the interchange above. Summing all the traces together there are a lot of cancellations and the final result is

$$8\left[\frac{k \cdot p}{k' \cdot p} + \frac{k' \cdot p}{k \cdot p} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \tag{185}$$

or, using that $k \cdot p = m\omega$ and $k' \cdot p = m\omega'$

$$8\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \tag{186}$$

Now let us return to the calculation of the probability, it can be written as

$$\begin{aligned} P = &\left|\frac{\pi e^2 (2\pi)^4 \delta^4 (p' + k' - p - k)}{2V^2 \sqrt{2EE'\omega\omega'}}\right|^2 8\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] = \\ VT &\frac{\pi^2 e^4}{V^4 EE'\omega\omega'} (2\pi)^4 \delta^4 (p' + k' - p - k) \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \end{aligned} \tag{187}$$

Calculating the probability per unit time and summing over inseparable final states we have

$$\begin{aligned} \frac{P}{T} &= \frac{\pi^2 e^4}{V^3 EE'\omega\omega'} \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \\ &\quad (2\pi)^4 \delta^4 (p' + k' - p - k) \frac{V}{(2\pi)^3} d^3\mathbf{p}' \frac{V}{(2\pi)^3} d^3\mathbf{k}' \end{aligned} \tag{188}$$

We get rid of the space-like part of the delta function if we perform one of the integrals, say the one over $\mathbf{p}'$, which leaves us with

$$\frac{P}{T} = \frac{\pi^2 e^4}{(2\pi)^2 V E E' \omega \omega'}\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \delta(E' + \omega' - E - \omega) d^3\mathbf{k}' \quad (189)$$

To get rid of the last delta function we can use that $\omega' = |\mathbf{k}'|$ to rewrite $d^3\mathbf{k}' = (\omega')^2 d\omega' d\Omega_k$. What complicates things slightly is that since we already used the space-like delta function, $E' = \sqrt{|\mathbf{p}'|^2 + m^2} = \sqrt{|\mathbf{k} - \mathbf{k}'|^2 + m^2} = \sqrt{\omega^2 + \omega'^2 - 2\omega\omega' \cos(\theta) + m^2}$ is also a function of ω'.

To perform the integration we have to write

$$\delta(E' + \omega' - E - \omega) d\omega' = \delta(E' + \omega' - E - \omega) \frac{d\omega'}{d(E' + \omega')} d(E' + \omega') = \frac{d\omega'}{d(E' + \omega')} \quad (190)$$

Now

$$\frac{d(E' + \omega')}{d\omega'} = \frac{\omega' - \omega \cos(\theta)}{E'} + 1 \quad (191)$$

which, using that $E' = \omega - \omega' + m$, we can write as

$$\frac{d(E' + \omega')}{d\omega'} = \frac{\omega(1 - \cos(\theta)) + m}{E'} = \frac{\omega\omega'(1 - \cos(\theta)) + m\omega'}{E'\omega'} = \frac{k \cdot k' + p \cdot k'}{E'\omega'} = \frac{(k' + p') \cdot k'}{E'\omega'} = \frac{p' \cdot k'}{E'\omega'} = \frac{p \cdot k}{E'\omega'} = \frac{E\omega}{E'\omega'} \quad (192)$$

where we have used that $p + k = p' + k'$ which in turn implies (by squaring) that $p \cdot k = p' \cdot k'$.

Finally, to get a number which is independent of how often we throw in photons (i.e. the cross section) we have to divide with the flux. Since the photon wave-function is normalized to one photon in the whole universe and since the speed of the photon is 1 (in natural units), the incoming flux is $\frac{1}{V}$. This gives us an expression for the cross section as

$$d\sigma = \frac{\pi^2 e^4}{(2\pi)^2 E E' \omega \omega'}\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] \frac{E'\omega'^3}{E\omega} d\Omega = \frac{e^4}{4m^2}\left(\frac{\omega'}{\omega}\right)^2 \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon' \cdot \varepsilon)^2 - 2\right] d\Omega \quad (193)$$

which is the famous Klein-Nishina formula for the cross section of Compton scattering.

13 The photon propagator

Remember that in a previous section we derived the electron/positron propagator which could be written as

$$G_{ab}(x_1 - x_2) = \langle 0|T\left[\psi_a(x_1)\bar{\psi}_b(x_2)\right]|0\rangle \tag{194}$$

As we will see in the next section, there is also a *photon* propagator corresponding to a photon propagating between two points x_1 and x_2. As one can guess, it can be written as

$$D_{\mu\nu}(x_1 - x_2) = \langle 0|T\left[A_\mu(x_1)A_\nu(x_2)\right]|0\rangle \tag{195}$$

Let us calculate it here so that we can use it when it appears in the next section. Inserting the expressions for the photon fields we have

$$\begin{aligned}
&\sum_{k_1,\alpha_1}\sum_{k_2,\alpha_2}\frac{4\pi}{2V\sqrt{\omega_1\omega_2}}\langle 0|T\Big[(a_1^\dagger\varepsilon_{1\mu}e^{ik_1\cdot x_1} + a_1\varepsilon_{1\mu}^\star e^{-ik_1\cdot x_1}) \\
&\qquad (a_2^\dagger\varepsilon_{2\nu}e^{ik_2\cdot x_2} + a_2\varepsilon_{2\nu}^\star e^{-ik_2\cdot x_2})\Big]|0\rangle = \\
&\sum_{k_1,\alpha_1}\sum_{k_2,\alpha_2}\frac{4\pi}{2V\sqrt{\omega_1\omega_2}}\Big\{\theta\left(t_1 - t_2\right)\varepsilon_{1\mu}^\star\varepsilon_{2\nu}e^{-ik_1\cdot x_1}e^{ik_2\cdot x_2}\langle 0|a_1a_2^\dagger|0\rangle \\
&\qquad \theta\left(t_2 - t_1\right)\varepsilon_{2\nu}^\star\varepsilon_{1\mu}e^{-ik_2\cdot x_2}e^{ik_1\cdot x_1}\langle 0|a_2a_1^\dagger|0\rangle\Big\}
\end{aligned} \tag{196}$$

Remembering that

$$\left[a_{k,\alpha}, a_{k',\alpha'}^\dagger\right] = \delta_{k,k'}\delta_{\alpha,\alpha'} \qquad \alpha = 1,2 \tag{197}$$

$$\tag{198}$$

we find

$$\begin{aligned}
D_{\mu\nu}(x_1 - x_2) = \sum_{\mathbf{k}}\frac{4\pi}{2V\omega}&\Big\{\theta\left(t_1 - t_2\right)\left(\left(\varepsilon_\mu^{(1)}\right)^\star\varepsilon_\nu^{(1)} + \left(\varepsilon_\mu^{(2)}\right)^\star\varepsilon_\nu^{(2)}\right)e^{-ik\cdot(x_1 - x_2)} \\
&\theta\left(t_2 - t_1\right)\left(\left(\varepsilon_\nu^{(1)}\right)^\star\varepsilon_\mu^{(1)} + \left(\varepsilon_\nu^{(2)}\right)^\star\varepsilon_\mu^{(2)}\right)e^{ik\cdot(x_1 - x_2)}\Big\}
\end{aligned} \tag{199}$$

This expression does not look too satisfying since the rather messy expression $\left(\left(\varepsilon_\nu^{(1)}\right)^\star \varepsilon_\mu^{(1)} + \left(\varepsilon_\nu^{(2)}\right)^\star \varepsilon_\mu^{(2)}\right)$ appears. In fact, using the explicit expressions for the polarization vectors

$$
\begin{aligned}
\varepsilon_k^{(0)} &= (1,0,0,0) \\
\varepsilon_k^{(1)} &= (0,\bar{\varepsilon}_1) \\
\varepsilon_k^{(2)} &= (0,\bar{\varepsilon}_2) \\
\varepsilon_k^{(3)} &= (0,\frac{\mathbf{k}}{|\mathbf{k}|})
\end{aligned}
\tag{200}
$$

where $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$ are two three dimensional unit vectors which are orthogonal both to $\mathbf{k}$ and to each other, one may show that

$$
-\left(\varepsilon_\nu^{(0)}\right)^\star \varepsilon_\mu^{(0)} + \left(\varepsilon_\nu^{(1)}\right)^\star \varepsilon_\mu^{(1)} + \left(\varepsilon_\nu^{(2)}\right)^\star \varepsilon_\mu^{(2)} + \left(\varepsilon_\nu^{(3)}\right)^\star \varepsilon_\mu^{(3)} = -g_{\mu\nu}. \tag{201}
$$

One way to prove it would be to say that it is a symmetric matrix with eigenvalues (-1,1,1,1), therefore it is diagonalizable to $-g_{\mu\nu}$ using an orthogonal matrix M so that $M^T(-\varepsilon_0\varepsilon_0^T + \varepsilon_1\varepsilon_1^T + \varepsilon_2\varepsilon_2^T + \varepsilon_3\varepsilon_3^T)M = -g$. But by multiplying on the left with M and on the right with M^T we find that $(-\varepsilon_0\varepsilon_0^T + \varepsilon_1\varepsilon_1^T + \varepsilon_2\varepsilon_2^T + \varepsilon_3\varepsilon_3^T) = -MgM^T = -g$.

Furthermore we may rewrite the $-\left(\varepsilon_\nu^{(0)}\right)^\star \varepsilon_\mu^{(0)} + \left(\varepsilon_\nu^{(3)}\right)^\star \varepsilon_\mu^{(3)}$ piece using the four vectors $k_\mu = (\omega, -\mathbf{k})$ and $a_\mu = (\omega, \mathbf{k})$ as

$$
\left(-\left(\varepsilon_\nu^{(0)}\right)^\star \varepsilon_\mu^{(0)} + \left(\varepsilon_\nu^{(3)}\right)^\star \varepsilon_\mu^{(3)}\right) = -\frac{1}{2\omega^2}\left(a_\mu k_\nu + k_\mu a_\nu\right). \tag{202}
$$

We thus find that we may write the propagator

$$
\begin{aligned}
D_{\mu\nu}(x_1 - x_2) = \sum_k \left(-g_{\mu\nu} + \frac{1}{2\omega^2}\left(a_\mu k_\nu + k_\mu a_\nu\right)\right) \times \\
\frac{4\pi}{2V\omega}\left\{\theta\left(t_1 - t_2\right) e^{-ik\cdot(x_1 - x_2)} + \theta\left(t_2 - t_1\right) e^{-ik\cdot(x_2 - x_1)}\right\}
\end{aligned}
\tag{203}
$$

Again using the fact that for a high density of states we can switch the sum over $\mathbf{k}$ for an integral over the density of states

$$
\sum_{\mathbf{k}} = V \int \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{204}
$$

and again using the rewriting of the energy dependence as

$$-e^{-i\omega|t|} = \frac{1}{2\pi i}\int_{-\infty}^{\infty} dk_0 \frac{e^{-ik_0 t}}{k_0^2 - |\mathbf{k}|^2 + i\varepsilon} \tag{205}$$

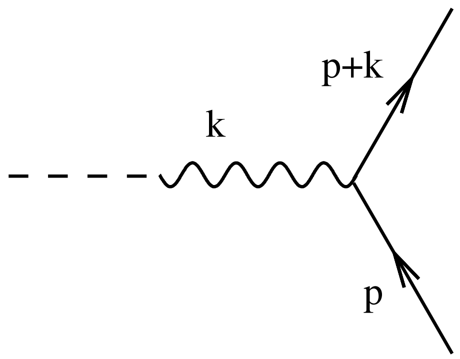
we can write the propagator as

$$\begin{aligned} D_{\mu\nu}(x_1 - x_2) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} 4\pi \left(g_{\mu\nu} - \frac{1}{2\omega^2}\left(a_\mu k_\nu + k_\mu a_\nu\right)\right) \times \\ \Big\{ \theta(t_1 - t_2) \frac{1}{2\pi i} \int \frac{dk_0}{k^2 + i\varepsilon} e^{ik_0(t_1-t_2)} e^{i\mathbf{k}\cdot(\mathbf{x_1}-\mathbf{x_2})} + \\ \theta(t_2 - t_1) \frac{1}{2\pi i} \int \frac{dk_0}{k^2 + i\varepsilon} e^{ik_0(t_1-t_2)} e^{-i\mathbf{k}\cdot(\mathbf{x_1}-\mathbf{x_2})} \Big\}. \end{aligned} \tag{206}$$

and changing the sign of the integration variable $\mathbf{k}$ in the second term, we see that both the term for $t_1 > t_2$ and the term for $t_2 > t_1$ have exactly the same form, just as for the electron/positron propagator, so we can write it in a compact form

$$D_{\mu\nu}(x_1 - x_2) = -i4\pi \int \frac{d^4k}{(2\pi)^4} \left(g_{\mu\nu} - \frac{1}{2\omega^2}\left(a_\mu k_\nu + k_\mu a_\nu\right)\right) \frac{e^{-ik\cdot(x_1-x_2)}}{k^2 + i\varepsilon}. \tag{207}$$

This is very nice except for the "$ak + ka$" piece which ruins the covariant form of the propagator. However, the fact that this piece contains explicit factors of k_μ which is *the same* four momentum as the momentum which flows in the propagator saves us. To see this, let us consider a part of a Feynman diagram looking like



The electron line corresponds to an expression (up to normalization constants)

$$\bar{u}(p+k)\gamma^\mu u(p), \tag{208}$$

To the γ^μ is connected the momentum space propagator $D_{\mu\nu}(k)$. The part of $D_{\mu\nu}$ proportional to k_μ together with (208) gives a contribution proportional to

$$\begin{aligned} \bar{u}(p+k)\not{k}u(p) = \bar{u}(p+k)\left(\not{k}+\not{p}-\not{p}\right)u(p) = \\ \{\text{Dirac equation}\} = \bar{u}(p+k)\,(m-m)\,u(p) = 0. \end{aligned} \tag{209}$$

Therefore the piece proportional to $k_\mu a_\nu$ does not give any contribution to physical processes and can consequently be dropped. Similarly, the piece proportional to $a_\mu k_\nu$ cancels in the other end of the propagator. The full proof of this fact is a little bit involved. One needs to check that it is true also when the photon propagator ends on an internal electron propagator and not on an external line as in the simple example above. If you believe me for now that this is true we can write the photon propagator simply as

$$D_{\mu\nu}(x_1-x_2) = -i4\pi g_{\mu\nu}\int\frac{d^4k}{(2\pi)^4}\frac{e^{-ik\cdot(x_1-x_2)}}{k^2+i\varepsilon}. \tag{210}$$

14 Electron-electron scattering

Let us now consider scattering of two electrons. Let us assume that initially they have momenta and spin p_1, s_1 and p_2, s_2. They are scattered into electrons with momenta and spin p_3, s_3 and $p_4.s_4$. We therefore take as the initial state

$$|\mathrm{i}\rangle = b_1^\dagger b_2^\dagger|0\rangle \tag{211}$$

and as the final state we take

$$|\mathrm{f}\rangle = b_4^\dagger b_3^\dagger|0\rangle \tag{212}$$

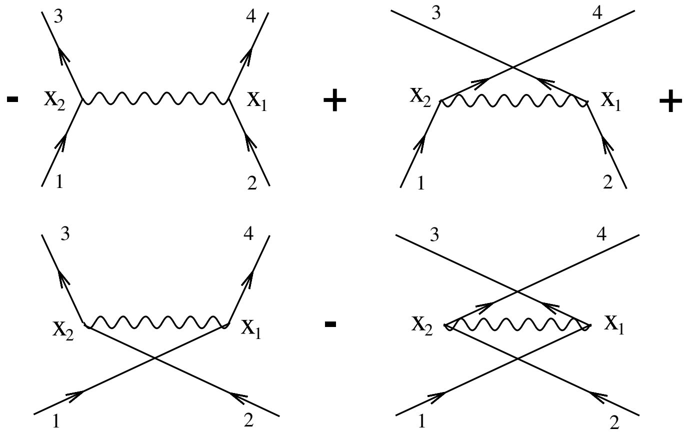
The probability amplitude is given by the usual expression

$$\mathcal{M} = \langle \mathrm{f}|T\ e^{-i\int H_I}|\mathrm{i}\rangle \tag{213}$$

and the first non-trivial term is the term which is second order in H_I. Since there are 4 b operators in the initial and final states the only combination which will give anything is when we choose the b operators from the 4 ψ fields. The expression is therefore

$$\begin{aligned}\langle 0|b_3 b_4 b^\dagger b' b''^\dagger b''' b_1^\dagger b_2^\dagger|0\rangle &= \\ -\delta_{1,p'''}\delta_{2,p'}\delta_{3,p''}\delta_{4,p} &+ \delta_{1,p'''}\delta_{2,p'}\delta_{3,p}\delta_{4,p''} \\ +\delta_{1,p'}\delta_{2,p'''}\delta_{3,p''}\delta_{4,p} &- \delta_{1,p'}\delta_{2,p'''}\delta_{3,p}\delta_{4,p''} \\ &+ \text{non}-\text{connected pieces} \qquad (214)\end{aligned}$$

where we have used a shorthand notation in that each delta function also comes with a corresponding delta for the spin dependence so that $\delta_{1,p}$ really means $\delta_{1,p}\delta_{\alpha_1,\alpha}$. These terms can be graphically represented as



Again using the trick of changing the integration variables x_1 and x_2 we see that the two last diagrams are equal to the two first diagrams which leaves us with only two expressions to be calculated. The one corresponding to the first diagram we can write as

$$\begin{aligned}&\frac{e^2}{4V^2\sqrt{E_1E_2E_3E_4}}(\bar{u}_4\gamma^\mu u_2)(\bar{u}_3\gamma^\nu u_1)\int d^4x_1\int d^4x_2 \\ &\quad e^{i(p_4-p_2)\cdot x_1}e^{i(p_3-p_1)\cdot x_2}\langle 0|T\left[A_\mu(x_1)A_\nu(x_2)\right]|0\rangle \qquad (215)\end{aligned}$$

where in the last expression we recognize the photon propagator $D_{\mu\nu}(x_1-x_2)$ which we calculated in the previous section. Inserting the expression we obtained we get

$$\frac{e^2}{4V^2\sqrt{E_1E_2E_3E_4}}(\bar{u}_4\gamma^\mu u_2)(\bar{u}_3\gamma^\nu u_1)\int d^4x_1\int d^4x_2\int\frac{d^4k}{(2\pi)^4}$$
$$e^{i(p_4-p_2-k)\cdot x_1}e^{i(p_3-p_1+k)\cdot x_2}\frac{4\pi g_{\mu\nu}}{i(k^2+i\varepsilon)} \tag{216}$$

and performing the x_1 and x_2 integrals we get

$$\frac{e^2}{4V^2\sqrt{E_1E_2E_3E_4}}(\bar{u}_4\gamma^\mu u_2)(\bar{u}_3\gamma^\nu u_1)\int\frac{d^4k}{(2\pi)^4}$$
$$(2\pi)^4\delta(p_4-p_2-k)(2\pi)^4\delta(p_3-p_1+k)\frac{4\pi g_{\mu\nu}}{i(k^2+i\varepsilon)} \tag{217}$$

We can get rid of one of the delta functions by doing the k integral

$$\frac{e^2}{4V^2\sqrt{E_1E_2E_3E_4}}(\bar{u}_4\gamma^\mu u_2)(\bar{u}_3\gamma^\nu u_1)$$
$$(2\pi)^4\delta(p_1+p_2-p_3-p_4)\frac{4\pi g_{\mu\nu}}{i((p_1-p_3)^2+i\varepsilon)} \tag{218}$$

Notice that the remaining delta function expresses the total conservation of momentum. The second diagram can be easily calculated when we realize that the only thing that differs between the second and the first diagram is that we have to switch 3 and 4 and also the overall sign. Totally we therefore have the probability amplitude

$$\mathcal{M}=\frac{e^2(2\pi)^4\delta(p_1+p_2-p_3-p_4)}{4V^2\sqrt{E_1E_2E_3E_4}}\left\{(\bar{u}_4\gamma^\mu u_2)\frac{4\pi g_{\mu\nu}}{i((p_1-p_3)^2+i\varepsilon)}(\bar{u}_3\gamma^\nu u_1)-\right.$$
$$\left.(\bar{u}_3\gamma^\mu u_2)\frac{4\pi g_{\mu\nu}}{i((p_1-p_4)^2+i\varepsilon)}(\bar{u}_4\gamma^\nu u_1)\right\} \tag{219}$$

If we assume that the incoming electrons are unpolarized so that we have to average over the incoming spins and that we do not observe the spin of the outgoing electrons so that we have to sum over the probabilities of observing different outgoing spin, we have to include a sum over

$$\frac{1}{2}\sum_{s_1}\frac{1}{2}\sum_{s_2}\sum_{s_3}\sum_{s_4} \tag{220}$$

Introducing the notation

$$\begin{aligned} t &= (p_1 - p_3)^2 = (p_4 - p_2)^2 \\ u &= (p_1 - p_4)^2 = (p_3 - p_2)^2 \end{aligned} \tag{221}$$

we can write the total probability as

$$\begin{aligned} P &= T\frac{e^4(2\pi)^6\delta(p_1+p_2-p_3-p_4)}{4V^3E_1E_2E_3E_4} \\ &\quad \frac{1}{4}\sum_{s_1,s_2,s_3,s_4}\left|\frac{(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma^\mu u_1)}{t} - \frac{(\bar{u}_3\gamma_\mu u_2)(\bar{u}_4\gamma^\mu u_1)}{u}\right|^2 \end{aligned} \tag{222}$$

which, using the expressions for the spin sums, can be written as

$$\begin{aligned} P &= T\frac{e^4(2\pi)^6\delta(p_1+p_2-p_3-p_4)}{4V^3E_1E_2E_3E_4} \\ &\quad \frac{1}{4}\left\{\frac{\mathrm{Tr}((\not{p}_4+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu)\mathrm{Tr}((\not{p}_3+m)\gamma^\mu(\not{p}_1+m)\gamma^\nu)}{t^2}\right. \\ &\quad +\frac{\mathrm{Tr}((\not{p}_3+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu)\mathrm{Tr}((\not{p}_4+m)\gamma^\mu(\not{p}_1+m)\gamma^\nu)}{u^2} \\ &\quad -\frac{\mathrm{Tr}((\not{p}_4+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_3+m)\gamma^\mu(\not{p}_1+m)\gamma^\nu)}{tu} \\ &\quad \left.-\frac{\mathrm{Tr}((\not{p}_3+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_4+m)\gamma^\mu(\not{p}_1+m)\gamma^\nu)}{tu}\right\} \end{aligned} \tag{223}$$

Let us explicitly calculate the first trace and leave the other ones as an exercise. To do this we observe that

$$\begin{aligned} \mathrm{Tr}((\not{p}+m)\gamma_\mu(\not{q}+m)\gamma_\nu) &= \mathrm{Tr}(\not{p}\gamma_\mu\not{q}\gamma_\nu) + m^2\mathrm{Tr}(\gamma_\mu\gamma_\nu) = \\ &\quad 4\left(p_\mu q_\nu - (p\cdot q)g_{\mu\nu} + p_\nu q_\mu + m^2 g_{\mu\nu}\right) \end{aligned} \tag{224}$$

giving the result for the first term

$$\begin{aligned} 16\Big[&2(p_1\cdot p_2)(p_3\cdot p_4) + 2(p_1\cdot p_4)(p_2\cdot p_3) + 2(m^2 - p_2\cdot p_4)(p_1\cdot p_3) + \\ &2(m^2 - p_1\cdot p_3)(p_2\cdot p_4) + 4(m^2 - p_1\cdot p_3)(m^2 - p_2\cdot p_4)\Big] \end{aligned} \tag{225}$$

and using that the relation $p_1 + p_2 = p_3 + p_4$ implies

$$\begin{aligned} p_1\cdot p_2 &= p_3\cdot p_4 \\ p_1\cdot p_3 &= p_2\cdot p_4 \qquad (226) \\ p_1\cdot p_4 &= p_2\cdot p_3 \qquad (227) \end{aligned}$$

we can write it as

$$32\left[(p_1\cdot p_2)^2+(p_1\cdot p_4)^2+2m^2(m^2-p_1\cdot p_3)\right] \tag{228}$$

Computing also the remaining traces we get the full answer

$$\begin{aligned} P &= T\frac{e^4(2\pi)^6\delta(p_1+p_2-p_3-p_4)}{4V^3E_1E_2E_3E_4} \\ & 8\left\{\frac{(p_1\cdot p_2)^2+(p_1\cdot p_4)^2+2m^2(m^2-p_1\cdot p_3)}{t^2}\right. \\ & +\frac{(p_1\cdot p_2)^2+(p_1\cdot p_3)^2+2m^2(m^2-p_1\cdot p_4)}{u^2} \\ & \left.-\frac{2(p_1\cdot p_2)(2m^2-p_1\cdot p_2)}{tu}\right\} \end{aligned} \tag{229}$$

Let us choose center of mass coordinates such that

$$\begin{aligned} p_1 &= (E,\mathbf{p}) \\ p_2 &= (E,-\mathbf{p}) \\ p_3 &= (E,\mathbf{p}') \\ p_4 &= (E,-\mathbf{p}') \end{aligned} \tag{230}$$

and that $\mathbf{p}\cdot\mathbf{p}'=|\mathbf{p}|^2\cos(\theta)$. Then we see that we can write

$$\begin{aligned} p_1\cdot p_2 &= m^2+2\,|\mathbf{p}|^2 \\ p_1\cdot p_3 &= m^2+2\,|\mathbf{p}|^2\sin^2\left(\frac{\theta}{2}\right) \\ p_1\cdot p_4 &= m^2+2\,|\mathbf{p}|^2\cos^2\left(\frac{\theta}{2}\right) \end{aligned} \tag{231}$$

To calculate the cross section we need to sum over the probabilities of observing final states which are close to each other in momentum space. This we do by including the factors

$$\frac{V}{(2\pi)^3}d^3\mathbf{p}_3\frac{V}{(2\pi)^3}d^3\mathbf{p}_4 \tag{232}$$

and dividing by the incoming flux $\frac{v_1+v_2}{V}$ where v_1 and v_2 is the speed of the 1 and 2 particle respectively. In the center of mass system the speed of the two particles are equal and can be expressed as $\frac{|\mathbf{p}|}{E}$ so that the flux is $\frac{2|\mathbf{p}|}{EV}$. The integrals over the momenta can be taken care of in the usual way

$$\delta^4(p_1 + p_2 - p_3 - p_4) d^3\mathbf{p}_4 d^3\mathbf{p}_3 = \delta(E_1 + E_2 - E_3 - E_4) |\mathbf{p}_3|^2 \, d\,|\mathbf{p}_3| \, d\Omega \tag{233}$$

and since in the center of mass system we have that $E_4 = E_3 = \sqrt{|\mathbf{p}_3|^2 + m^2}$ we can write

$$\delta(E_1 + E_2 - 2E_3) |\mathbf{p}_3|^2 \frac{d\,|\mathbf{p}_3|}{d(2E_3)} d(2E_3) d\Omega = \frac{|\mathbf{p}_3| E_3}{2} d\Omega \tag{234}$$

Putting everything together we have

$$\begin{aligned}
d\sigma &= \frac{e^4(2\pi)^6}{4V^3E^4} \frac{V^2}{(2\pi)^6} \frac{EV}{2\,|\mathbf{p}|} \frac{|\mathbf{p}|\, E}{2} \\
&\quad 8\left\{ \frac{(p_1\cdot p_2)^2 + (p_1\cdot p_4)^2 + 2m^2(m^2 - p_1\cdot p_3)}{t^2} \right. \\
&\quad + \frac{(p_1\cdot p_2)^2 + (p_1\cdot p_3)^2 + 2m^2(m^2 - p_1\cdot p_4)}{u^2} \\
&\quad \left. - \frac{2(p_1\cdot p_2)(2m^2 - p_1\cdot p_2)}{tu} \right\} d\Omega \\
&= \frac{e^4}{2E^2} \left\{ \frac{(p_1\cdot p_2)^2 + (p_1\cdot p_4)^2 + 2m^2(m^2 - p_1\cdot p_3)}{t^2} \right. \\
&\quad + \frac{(p_1\cdot p_2)^2 + (p_1\cdot p_3)^2 + 2m^2(m^2 - p_1\cdot p_4)}{u^2} \\
&\quad \left. - \frac{2(p_1\cdot p_2)(2m^2 - p_1\cdot p_2)}{tu} \right\} d\Omega
\end{aligned} \tag{235}$$

This cross section simplifies in the non-relativistic and the ultra-relativistic cases. In the non-relativistic case we have that $|\mathbf{p}| \ll m$ so that to lowest order

$$\begin{aligned}
p_1\cdot p_2 &\approx m^2 \\
p_1\cdot p_3 &\approx m^2 \\
p_1\cdot p_4 &\approx m^2 \\
t &= (p_1 - p_3)^2 = 2m^2 - 2\left(m^2 + 2\,|\mathbf{p}|^2 \sin^2\left(\frac{\theta}{2}\right)\right)
\end{aligned} \tag{236}$$

$$
\begin{aligned}
&\approx -4\,|\mathbf{p}|^2 \sin^2\left(\frac{\theta}{2}\right) \\
u &= 2m^2 - 2\left(m^2 + 2\,|\mathbf{p}|^2 \cos^2\left(\frac{\theta}{2}\right)\right) \\
&\approx -4\,|\mathbf{p}|^2 \cos^2\left(\frac{\theta}{2}\right)
\end{aligned} \tag{237}
$$

which gives the expression

$$
d\sigma = \frac{e^4}{m^2 v^4}\left(\frac{1}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{1}{\cos^4\left(\frac{\theta}{2}\right)} - \frac{1}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)}\right) d\Omega \tag{238}
$$

where v is the velocity $v = \frac{|\mathbf{p}|}{m}$. Notice that the first term gives exactly the Rutherford cross-section. The two additional terms are of quantum mechanical origin and comes about because the particles that scatter are quantum mechanically *identical*. This means that the cross section has to be invariant under $\theta \to \pi - \theta$. The second term alone would be enough to achieve that. The third term is there however because scattering of identical Fermi particles is very much suppressed at $\theta = \frac{\pi}{2}$. This is essentially an effect of the Pauli principle which tells us that the total wave function has to be anti-symmetric.

In the ultra-relativistic limit, $|p| \gg m$, the cross section can similarly be written in a simple form

$$
d\sigma = \frac{e^4}{8E^2}\left(\frac{1+\sin^4\left(\frac{\theta}{2}\right)}{\cos^4\left(\frac{\theta}{2}\right)} + \frac{1+\cos^4\left(\frac{\theta}{2}\right)}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{2}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)}\right) d\Omega \tag{239}
$$

15 The Ward-Takahashi identity