Université de Bourgogne - Franche Comté

Supervised research work

Bachelor's degree in Fundamental Physics

Quantum mechanics: Demonstrating the Dirac equation

Rodolphe MOMIER
UBFC

Advisor: Pr. Claude LEROY
ICB-Dijon

January 14-18th 2019

Contents

1 Introduction

This work aims at giving the main lines of a demonstration of the **Dirac equation**, ruling the behavior of a charged particle under the influence of an EM field, and taking into account the **relativistic effects**. Starting with a little reminder about **Lagrange** and **Hamilton** equations of motion, we will progress through the evolution of modern physics, as chronologically as possible, starting from the **Klein-Gordon** equation. The **Dirac hamiltonian** will be redemonstrated and we will finish by developing the Dirac equation. Of course, all the calculations won't be developed but all the demonstrations have been done.

NB: in all this work, vectors will be denoted in **bold letters**.

2 Motion of a particle in an EM field

This part aims at demonstrating the Hamiltonian (and Lagrangian) of a particle under the influence of an EM field. The Hamiltonian formalism is the one used for quantum mechanics. We start by reminding the Lagrange and Hamilton equations of motion, then we write the Lagrangian which is needed to calculate the Hamiltonian. We can refer to [1] and [3] for further explanations.

2.1 Lagrange and Hamilton equations

For a physical system described by the generalized coordinates q_i and their derivatives $\dot{q}_i$ ($i = 1, ..., n$), the Lagrangian is

$$L = E_{\text{kin}} - E_{\text{pot}} = T - U \tag{1}$$

and the Lagrange equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad (i = \mathit{1}, ..., n)\,. \tag{2}$$

The Hamiltonian H is the **Legendre transform** of the Lagrangian:

$$H = \sum_{i=1}^{n} p_i \dot{q}_i - L \tag{3}$$

where $p_i = \dfrac{\partial L}{\partial \dot{q}_i}$ is the **generalized momentum**. We can demonstrate that the **Hamilton equations of motion** are:

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \qquad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \qquad (i = \mathit{1}, ..., n)\,.$$

2.2 Lagrangian and Hamiltonian of a particle of a particle in an EM field

We consider a particle of mass m, of charge q, of speed $\mathbf{v}$ in an EM field $(\mathbf{E}, \mathbf{B})$. The only force applied on the particle is

$$\mathbf{F}_{\text{Lorentz}} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})\,. \tag{4}$$

The EM field $(\mathbf{E}, \mathbf{B})$ can be associated to a gauge $(\mathbf{A}, V)$ defined by the relations:

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla} V \quad \text{and} \quad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}\,. \tag{5}$$

The potential energy U a priori only depends on $\boldsymbol{q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$ and $\dot{\boldsymbol{q}}$ so we write

$$L = \frac{1}{2} m \dot{\boldsymbol{q}}^2 - U(\boldsymbol{q}, \dot{\boldsymbol{q}})\,. \tag{6}$$

Obviously for a particle in an EM field, the generalized coordinates q_i $(i = 1, ..., 3)$ are the standard cartesian coordinates (x, y, z). Thus applying eq.(2) on eq.(6) it gives -for instance- for the x coordinate

$$m\ddot{x} = -\frac{\partial U}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U}{\partial \dot{x}}\,. \tag{7}$$

We rewrite the Lorentz force using eq.(5):

$$\mathbf{F}_{\text{Lorentz}} = q\left[-\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla} V + \mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A})\right]\,.$$

Using eq.(44) in Appendix 7.1: $(\mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A}))_x = \frac{\partial}{\partial x}(\mathbf{A} \cdot \mathbf{v}) - \mathbf{v} \cdot \boldsymbol{\nabla} A_x$ we can express F_x the projection along (Ox) of $\mathbf{F}_{\text{Lorentz}}$ as

$$F_x = q\left[-\frac{\partial A_x}{\partial x} - \frac{\partial V}{\partial x} + \frac{\partial}{\partial x}(\mathbf{A} \cdot \mathbf{v}) - \mathbf{v} \cdot \boldsymbol{\nabla} A_x\right]\,.$$

We note that $A_x = \frac{\partial}{\partial \dot{x}}(\mathbf{A} \cdot \mathbf{v})$ thus

$$m\ddot{x} = -\frac{\partial}{\partial x} q(V - \mathbf{A} \cdot \mathbf{v}) - q\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}(\mathbf{A} \cdot \mathbf{v} - V)\,. \tag{8}$$

By comparing eq.(7) with eq.(8) we deduce $U = q(V - \mathbf{A} \cdot \mathbf{v})$ hence the Lagrangian

$$L = \frac{1}{2} m \mathbf{v}^2 - qV + q\mathbf{A} \cdot \mathbf{v} \Leftrightarrow \boxed{L = \frac{1}{2} m \mathbf{v}^2 + q(\mathbf{v} \cdot \mathbf{A} - V)}\,.$$

With $p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$ $(\alpha = x, y, z)$ the **generalized momentum**, we obtain

$$\mathbf{p} = \boldsymbol{\pi} + q\mathbf{A} \tag{9}$$

where $\boldsymbol{\pi}$ is the **momentum**. Using eq.(3) and eq.(9) and rearranging the terms, we can write

$$\boxed{H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + qV}\,.$$

2.3 Discussion

The transition from classical to quantum mechanics imposes

$$\mathbf{p} = -i\hbar\boldsymbol{\nabla} \qquad \text{and} \qquad E = i\hbar\frac{\partial}{\partial t}\,. \tag{10}$$

We also notice that the generalized momentum becomes $\mathbf{p} - q\mathbf{A}$ in the presence of an EM field. For the rest of this work, we will consider

$$\mathbf{p} \to \mathbf{p} - q\mathbf{A} \qquad \text{and} \qquad E \to E - qV\,. \tag{11}$$

Consequently, the Hamiltonian of a particle under the influence of a scalar potential V and vector potential $\mathbf{A}$ is

$$\boxed{H = \frac{1}{2m}(-i\hbar\boldsymbol{\nabla} - q\mathbf{A})^2 + qV}\,.$$

3 Klein-Gordon equation

3.1 Introducing the Klein-Gordon equation

This part is based on the work that has been done in [2] with elements from [5]. In 1926, W. Gordon and O. Klein adapted the Schrödinger equation to take into account the relativistic effects. Starting from the **relativistic energy** of a free particle

$$E^2 = p^2c^2 + m^2c^4\,, \tag{12}$$

one can rewrite eq.(12) as

$$-\hbar^2\frac{\partial^2}{\partial t^2} = -\hbar^2c^2 \bigtriangleup + m^2c^4 \tag{13}$$

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \bigtriangleup + \frac{m^2c^2}{\hbar^2}\right)\Psi = 0\,. \tag{14}$$

By denoting $\Box$ the **d'Alembertian** operator and Ψ being a wave function, we obtain the **Klein-Gordon equation**

$$\boxed{\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\Psi = 0}\,. \tag{15}$$

A more compact form can be found in the literature, using the covariant and contravariant derivatives $\partial_\mu = \left(\frac{\partial}{\partial t};\boldsymbol{\nabla}\right)$, $\partial^\mu = \left(\frac{\partial}{\partial t};-\boldsymbol{\nabla}\right)$ and the natural units $\hbar = c = 1$, K-G equation takes the final form

$$\boxed{\left(\partial_\mu\partial^\mu + m^2\right)\Psi = 0}\,. \tag{16}$$

K-G equation can be generalized to the case where the particle is under the influence of an EM field. By injecting eq.(11) and eq.(10) into eq.(12), we get

$$\boxed{\left[\left(i\hbar\frac{\partial}{\partial t} - qV\right)^2 - c^2\left(-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right)^2\right]\Psi = m^2c^4\Psi}\,.$$

3.2 Which problems did it cause ?

- First problem:

Equation (16) (or eq.(14),eq.(15)) is a **partial differential equation** of order 2 with respect to the time t. This equation can be solved with two initial conditions but unfortunately we do not have them. We can demonstrate that if we "split Ψ into two parts":

$$\Psi_A = \Psi + \frac{i\hbar}{mc^2}\frac{\partial\Psi}{\partial t} \qquad \text{and} \qquad \Psi_B = \Psi - \frac{i\hbar}{mc^2}\frac{\partial\Psi}{\partial t}$$

which implies

$$\Psi = \frac{1}{2}(\Psi_A + \Psi_B) \qquad \text{and} \qquad \frac{\partial\Psi}{\partial t} = \frac{mc^2}{2i\hbar}(\Psi_A - \Psi_B)\,,$$

we can rewrite (**linearize**) K-G equation on a matrix form:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_A\\ \Psi_B\end{pmatrix} = \begin{pmatrix} mc^2 - \frac{\hbar^2}{2m}\bigtriangleup & -\frac{\hbar^2}{2m}\bigtriangleup \\ \frac{\hbar^2}{2m}\bigtriangleup & mc^2 + \frac{\hbar^2}{2m}\bigtriangleup\end{pmatrix}\begin{pmatrix}\Psi_A\\ \Psi_B\end{pmatrix}. \tag{17}$$

Even if the second order differential equation has been converted into a system of first order differential equations, we still need **two initial conditions** to solve them.

- Second problem:

Let's consider $\mathbf{j}$ (current of probability) and ρ (density of probability)[1]

$$\mathbf{j} = \frac{\hbar}{2im}(\Psi^* \boldsymbol{\nabla} \Psi - \Psi \boldsymbol{\nabla} \Psi^*) \tag{18}$$

$$\rho = \frac{i\hbar}{2mc^2}(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t})\,. \tag{19}$$

Using eq.(18) and its time derivative $\dfrac{\partial \mathbf{j}}{\partial t}$ and using eq.(19), we can write a linear combination of the Klein-Gordon equation

$$\Psi^* \left(\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} - \triangle \Psi + \frac{m^2 c^2}{\hbar^2} \Psi \right) - \Psi \left(\frac{1}{c^2} \frac{\partial^2 \Psi^*}{\partial t^2} - \triangle \Psi^* + \frac{m^2 c^2}{\hbar^2} \Psi^* \right) = 0\,. \tag{20}$$

We develop eq.(20) taking into account the definitions eq.(18) and eq.(19) to obtain a **conservation equation**

$$\boxed{\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j} = 0}\,.$$

However, ρ (real by construction) can be **negative** which is contradictory to the definition of **density**[2].

3.3 Discussion

1. If we want a relativistic equation, we need a **symmetry** between time and space coordinates.
2. We want a **first order differential equation** (Schrödinger equation works nicely for the non-relativistic cases, proving that first order equations are enough to describe the particles).
3. Examining eq.(17), we can see that the space derivatives are of second order due to the Laplacian. We could apply the same "trick" that we use for harmonic oscillators [3]. This leads us to a conclusion: the dimension of the matrix will double, which means Ψ_A and Ψ_B are both components of a 2-dimensional vector space, hence Ψ is a **4-component** wave function called **Spinor**. For the rest of this work, we will denote $\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$ and $\rho(\mathbf{r}, t) = \sum_{i=1}^{4} |\Psi_i(\mathbf{r}, t)|^2$.
4. We notice from eq.(12) that energy reduces to mc^2 when $v = 0$. We search the simplest equation being **linear** in $i\hbar\frac{\partial}{\partial t}$ and $\mathbf{p} = -i\hbar\boldsymbol{\nabla}$, which leads to us towards trying an equation of the type

$$\boxed{i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = (c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2)\Psi(\mathbf{r}, t) = H_{\text{Dirac}} \Psi(\mathbf{r}, t)}\,. \tag{21}$$

H_{Dirac} denotes the **Dirac Hamiltonian** that we will determine in the next part.

[1] for free Spin-0 particles

[2] See [2] p.971

[3] A trivial example would be an harmonic oscillator ruled by the equation $m\ddot{x} = -kx$. Using the Hamilton equations of motion, we write:

$$\dot{x} = \frac{p}{m} \qquad \dot{p} = -kx$$

This can be rewritten using matrices, hence:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix}$$

which constitutes a system of first order differential equations.

4 Dirac equation

4.1 Preliminary discussion

1. Equation (21) is the correct general form of the equation we want. If we observe eq.(21), we can deduce that β has to be a matrix. Indeed, $\Psi(\mathbf{r}, t)$ has 4 components and mc^2 is a scalar. To be even more precise, βmc^2 has to be homogenous to an energy, meaning β is a 4×4 matrix with **no physical dimension**.
2. Consequently, $c\boldsymbol{\alpha} \cdot \mathbf{p} = c(\alpha_x p_x + \alpha_y p_y + \alpha_z p_z)$ has to be a 4×4 matrix too, meaning $\alpha_{x,y,z}$ are 4×4 matrices.
3. In the case of a static force field with $V(r)$ being a potential, eq.(21) becomes

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = (c\boldsymbol{\alpha} \cdot \mathbf{p} + V(r)\mathbb{1}_4 + \beta mc^2)\Psi(\mathbf{r}, t) = H_{\text{Dirac}}\Psi(\mathbf{r}, t)\,.$$

4.2 Determination of H_{Dirac}

By denoting $\boldsymbol{\alpha} \cdot \boldsymbol{\nabla} = \sum_{\eta=x,y,z} \alpha_\eta \dfrac{\partial}{\partial \eta}$, eq.(21) can be rewritten as

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \frac{imc}{\hbar}\beta + \sum_{\eta=x,y,z} \alpha_\eta \frac{\partial}{\partial \eta}\right)\Psi(\mathbf{r}, t) = 0\,. \tag{22}$$

We have to determine $\boldsymbol{\alpha}$ and $\mathbf{p}$. In order to get rid of the crossed derivative terms $\left(\frac{\partial}{\partial \tau} \cdot \frac{\partial}{\partial \eta}\right)$, we define a new operator

$$\left(\frac{1}{c}\frac{\partial}{\partial t} - \frac{imc}{\hbar}\beta - \sum_{\tau=x,y,z} \alpha_\tau \frac{\partial}{\partial \tau}\right)$$

that we apply on eq.(22). Keeping in mind the Schwarz theorem and also that $\alpha_{x,y,z}$ don't necessarily commute between them, we obtain

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{1}{2}\sum_{\tau,\eta}(\alpha_\tau\alpha_\eta + \alpha_\eta\alpha_\tau)\frac{\partial^2}{\partial\eta\partial\tau} - \frac{imc}{\hbar}\sum_{\tau}(\alpha_\tau\beta + \beta\alpha_\tau)\frac{\partial}{\partial\tau} + \left(\frac{mc}{\hbar}\right)^2\beta^2\right)\Psi = 0\,. \tag{23}$$

However, eq.(23) has somehow been obtained from eq.(13) and eq.(14), thus eq.(23) has to be reduced to eq.(14) ie:

1. $\alpha_\tau\alpha_\eta + \alpha_\eta\alpha_\tau = 2\delta_{\tau\eta}$
2. $\beta^2 = \mathbb{1}_4$
3. $\alpha_\tau\beta + \beta\alpha_\tau = 0\,.$

These conditions are enough to determine one solution[4]

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathbb{O}_2 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbb{O}_2 \end{pmatrix} \qquad \beta = \begin{pmatrix} \mathbb{1}_2 & \mathbb{O}_2 \\ \mathbb{O}_2 & -\mathbb{1}_2 \end{pmatrix} \qquad \mathbb{O}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

involving the elements $\sigma_x, \sigma_y, \sigma_z$ called matrices of **Pauli**. These matrices are defined by

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

Actually, the **spin** is a **linear combination of the matrices of Pauli**. We obtain for H_{Dirac} the following matrix:

$$\boxed{H_{\text{Dirac}} = \begin{pmatrix} mc^2\mathbb{1}_2 & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & -mc^2\mathbb{1}_2 \end{pmatrix}}\,.$$

[4] $\boldsymbol{\alpha}$ means $\alpha_x = \begin{pmatrix} \mathbb{O}_2 & \sigma_x \\ \sigma_x & \mathbb{O}_2 \end{pmatrix}$ etc

4.3 More about the Dirac Hamiltonian

The most important thing to see is that the **spin** appears naturally in Dirac's theory. We denote

$$\mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \qquad \mathbf{J} = \mathbf{L} + \frac{\hbar}{2}\mathbf{\Sigma}\,.$$

$V(r)$ obviously commutes with α and β (and any other matrix). It also commutes with $\mathbf{L}$ because $V(r)$ is of spherical symmetry, hence

$$[\mathbf{J}, H_{\text{Dirac}}] = [\mathbf{J}, c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2]\,.$$

With eq.(46) of Appendix 7.2, we write

$$[\mathbf{L}, H_{\text{Dirac}}] = i\hbar c(\boldsymbol{\alpha}\times\mathbf{p}) \qquad [\mathbf{\Sigma}, H_{\text{Dirac}}] = 2ic(\mathbf{p}\times\boldsymbol{\alpha})$$

and we obtain

$$[\mathbf{J}, H_{\text{Dirac}}] = 0\,.$$

4.4 Discussion

$\frac{\hbar}{2}\mathbf{\Sigma}$ is the **momentum of Spin** of the particle, often denoted $\mathbf{S}$ in the litterature, because when added to the orbital (angular) momentum $\mathbf{L}$, the resulting angular momentum $\mathbf{J}$ commutes with the Hamiltonian ie. $\mathbf{J}$ is a **constant of motion**.

4.5 Developing the equation of Dirac

Taking into account $E \to E - qV$, $\mathbf{p} \to \mathbf{p} - q\mathbf{A}$ and in the **weak relativistic limit**, we can write:

$$\left[E - qV - c\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A}) - \beta mc^2\right]\Psi = 0\,,$$

which becomes by multiplying the LHS by $\left[E - qV + c\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A}) + \beta mc^2\right]$:

$$\left[(E - qV)^2 - c^2[\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A})]^2 - \overbrace{\beta^2}^{=\mathbb{1}_4} m^2c^4\right]\Psi \tag{24}$$

$$= \left[c(E - qV)\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A}) - c\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A})(E - qV) + \cancel{(E - qV)\beta mc^2}\right]\Psi \tag{25}$$

$$+ \left[-\cancel{\beta mc^2(E - qV)} + \underbrace{c\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A})\beta mc^2 + \beta mc^2 c\boldsymbol{\alpha}(\mathbf{p} - q\mathbf{A})}_{0}\right]\Psi = 0 \tag{26}$$

We demonstrate with eq.(47) of Appendix 7.3 that

$$[\boldsymbol{\alpha}\cdot(\mathbf{p} - q\mathbf{A})]^2 = (\mathbf{p} - q\mathbf{A})^2 - q\hbar\mathbf{\Sigma}\cdot\mathbf{B}\,, \tag{27}$$

So eq.(24) can be written as

$$[(E - qV)^2 - c^2(\mathbf{p} - q\mathbf{A})^2 + q\hbar c^2\mathbf{\Sigma}\cdot\mathbf{B} - m^2c^4]\Psi = 0\,.$$

Knowing that $\boldsymbol{\alpha}$ doesn't depend on time, we can develop eq.(25) and it becomes

$$(i\hbar qc\boldsymbol{\alpha}\cdot\mathbf{E})\Psi\,. \tag{28}$$

With eq.(27) and eq.(28), the Dirac equation becomes

$$\boxed{\left[(E - qV)^2 - c^2(\mathbf{p} - q\mathbf{A})^2 + q\hbar c^2\mathbf{\Sigma}\cdot\mathbf{B} - m^2c^4 - i\hbar qc\boldsymbol{\alpha}\cdot\mathbf{E}\right]\Psi = 0}\,. \tag{29}$$

Equation (29) is called the **relativistic form** (including the EM field $\mathbf{E}, \mathbf{B}$) of the **Schrödinger equation**. We will get deeper into this equation in the next and last part.

5 A bit more explanation about the Dirac equation

5.1 No EM field

Our starting point is the follwing equation:

$$\boxed{\underbrace{\left(c\boldsymbol{\alpha}\cdot\mathbf{p}+\beta mc^2\right)}_{H_{\text{Dirac}}}\Psi = E\Psi} \,. \tag{30}$$

Using the solutions that we found in part. 4.2, H_{Dirac} (Dirac Hamiltonian) can be rewritten as

$$H_{\text{Dirac}} = \begin{pmatrix} mc^2 & 0 & p_z & p_x - ip_y \\ 0 & mc^2 & p_x + ip_y & -p_z \\ p_z & p_x - ip_y & -mc^2 & 0 \\ p_x + ip_y & -p_z & 0 & -mc^2 \end{pmatrix}$$

and eq.(30) becomes a system of 4 equations

$$\begin{cases} c(p_x - ip_y)\Psi_4 + cp_z\Psi_3 + (mc^2 - E)\Psi_1 = 0 \\ c(p_x + ip_y)\Psi_3 - cp_z\Psi_4 + (mc^2 - E)\Psi_2 = 0 \\ c(p_x - ip_y)\Psi_2 + cp_z\Psi_1 - (mc^2 + E)\Psi_3 = 0 \\ c(p_x + ip_y)\Psi_1 - cp_z\Psi_2 - (mc^2 + E)\Psi_4 = 0 \,, \end{cases}$$

or equivalently, a system of two equations

$$\boxed{\begin{cases} c\boldsymbol{\sigma}\cdot\mathbf{p}\Psi_v + (mc^2 - E)\Psi_u = 0 \\ c\boldsymbol{\sigma}\cdot\mathbf{p}\Psi_u + (mc^2 - E)\Psi_v = 0 \end{cases}} \,.$$

We obtain two useful relations

$$\Psi_u = \frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E - mc^2}\Psi_v \tag{31}$$

$$\Psi_v = \frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E + mc^2}\Psi_u \,, \tag{32}$$

hence

$$\boxed{\Psi_u = \frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{(E - mc^2)}\frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{(E + mc^2)}\Psi_u = \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p})}{\mathbf{p}\cdot\mathbf{p}}\Psi_v} \,. \tag{33}$$

It is obvious, starting from eq.(33), that $|\boldsymbol{\sigma}\cdot\mathbf{p}| = |\mathbf{p}| \simeq mv$ in the **non-relativistic** case (ie. γ (Lorentz factor) $\simeq 1$). We can also notice (with a Taylor expansion) that $E = mc^2\sqrt{1 + \frac{p^2c^2}{m^2c^4}} \simeq mc^2\left(1 + \frac{p^2c^2}{2m^2c^4}\right)$ thus

$$\boxed{E \simeq \frac{p^2}{2m} = \frac{mv^2}{2}}$$

which leads us to

$$\frac{\Psi_u}{\Psi_v} = \frac{2c}{v} \quad \text{ie.} \quad \boxed{\Psi_u >> \Psi_v} \quad \text{for} \quad \gamma \simeq 1 \,. \tag{34}$$

Equation (34) explains why Ψ_u is called the **large component** of the wave function. This component is the one involved in the Schrödinger equation when there is no potential energy and EM field. Considering $E' = E - mc^2$ and using relation eq.(47), we obtain from eq.(33) a well known **eigenvalue** equation (valid for Ψ_u and Ψ_v so for Ψ in general)

$$\boxed{E'\Psi = \frac{-\hbar^2}{2m}\nabla^2\Psi} \,. \tag{35}$$

Equation (35) is the **time-independant Schrödinger equation** for a free particle without an electromagnetic field.

5.2 In presence of an EM field

We start from eq.(30) taking into account eq.(11)

$$\boxed{[\boldsymbol{\alpha}\cdot(c\mathbf{p}-q\mathbf{A})+\beta mc^2+qV]\Psi=E\Psi},$$

which can be rewritten in the **two-component form**

$$\begin{cases}\boldsymbol{\sigma}\cdot(c\mathbf{p}-q\mathbf{A})\Psi_v+(mc^2+qV)\Psi_u=E\Psi_u\\ \boldsymbol{\sigma}\cdot(c\mathbf{p}-q\mathbf{A})\Psi_u-(mc^2-qV)\Psi_v=E\Psi_v\,.\end{cases}$$

If we focus on the large component (as before) we can write

$$(E-mc^2-qV)\Psi_u=\frac{\boldsymbol{\sigma}\cdot(c\mathbf{p}-q\mathbf{A})\boldsymbol{\sigma}\cdot(c\mathbf{p}-q\mathbf{A})}{E+mc^2-qV}\Psi_u\,,$$

or more condensed (with $E'=E-mc^2$):

$$\boxed{(E'-qV)\Psi_u=\frac{1}{2m}(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})K(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})\Psi_u}\,. \tag{36}$$

We notice the introduction of two new elements in eq.(36):

$$K=\frac{1}{1+\dfrac{E'-qV}{2mc^2}}\simeq\frac{1}{1+\varepsilon}\qquad\text{and}\qquad\boldsymbol{\Pi}=\mathbf{p}-\frac{q}{c}\mathbf{A}\,.$$

Equation (36) is, like the time-independant Schrödinger equation, an eigenvalue equation. This equation is very difficult to solve as it is. We will continue by trying to define some approximations, valid at the order $\frac{v^2}{c^2}$. First, we have to rewrite $(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})$.

With relation eq.(47) we can write

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})^2&=\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2-\frac{iq}{c}\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{A}+\mathbf{A}\times\mathbf{p})\\&=\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2-\frac{iq}{c}\boldsymbol{\sigma}\cdot(-i\hbar\boldsymbol{\nabla}\times\mathbf{A})\\&=\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2-\frac{q\hbar}{c}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\,.\end{aligned} \tag{37}$$

The factor K can be developped using the Taylor expansion of the expression $\frac{1}{1+\varepsilon}$. For the second order, we obtain

$$K\simeq1-\frac{E'-qV}{2mc^2}+o(\varepsilon^2)\,,$$

and eq.(36) becomes at the **first and second orders of approximation**:

$$\boxed{(E'-qV)\Psi_u=\left[\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}A\right)^2-\frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\right]\Psi_u}\ \text{ when }\ K=1$$

$$\boxed{(E'-qV)\Psi_u=\left[\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}A\right)^2-\frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})-\underbrace{(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})\frac{E'-qV}{4m^2c^2}(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})}_{*}\right]\Psi_u}\,.$$

We find in the literature that we cannot transform (*) as we did with eq.(36). R. Feynman[5] and J.J Sakurai[6] solved this problem but used quite complicated methods. We prefer using another method, easier to understood that will be partially described here, starting from the matrix equation

$$\Psi=\Omega\Psi_u\,,$$

[5] Feynman, R. P. Quantum electrodynamics. Reading: Addison-Wesley. (1961)
[6] Sakurai, J. J. Advanced quantum mechanics. Reading: Addison-Wesley. (1967)

where we denote

$$\Omega = 1 + \frac{(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2}{8m^2c^2} \qquad \text{and} \qquad \Omega^{-1} = 1 - \frac{(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2}{8m^2c^2}\,.$$

We quickly notice that $\Omega\Omega^{-1}$ is of order $\dfrac{v^4}{c^4}$ hence Ω^{-1} in the inverse transformation of Ω **only** at the order $\dfrac{v^2}{c^2}$. We multiply the LHS of eq.(37) by Ω^{-1}. With $\Psi = \Omega\Psi_u$ we get

$$\boxed{\Omega^{-1}(E' - qV)\Omega^{-1}\Psi = \frac{1}{2m}\Omega^{-1}\left[(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2 - (\boldsymbol{\sigma} \cdot \mathbf{\Pi})\frac{E' - qV}{2mc^2}(\boldsymbol{\sigma} \cdot \mathbf{\Pi})\right]\Omega^{-1}\Psi}\,. \tag{38}$$

The next step is to develop with the expressions of Ω and Ω^{-1}, but this would be too long. The main thing to remember is that we limit ourselves to an approximation at the order $\frac{v^2}{c^2}$, thus the terms of higher orders are ignored. Some of the **most important developments** that are used are listed below:

1.
$$(\boldsymbol{\sigma} \cdot \mathbf{\Pi})(E' - qV) - (E' - qV)(\boldsymbol{\sigma} \cdot \mathbf{\Pi}) = i\hbar q\boldsymbol{\sigma} \cdot (\boldsymbol{\nabla} V)\,, \tag{39}$$

2.
$$\boldsymbol{\sigma} \cdot \mathbf{\Pi} = \begin{pmatrix} p_z - \frac{q}{c}A_z & p_x - \frac{q}{c}A_x - ip_y + i\frac{q}{c}A_y \\ p_x - \frac{q}{c}A_x + ip_y - \frac{q}{c}A_y & -p_z + \frac{q}{c}A_z \end{pmatrix},$$

3.
$$\boldsymbol{\sigma} \cdot (\boldsymbol{\nabla} V) = \begin{pmatrix} \frac{\partial V}{\partial z} & \frac{\partial V}{\partial x} - i\frac{\partial V}{\partial y} \\ \frac{\partial V}{\partial x} + i\frac{\partial V}{\partial y} & -\frac{\partial V}{\partial z} \end{pmatrix}. \tag{40}$$

Applying operator (40) on a function φ, we can express that

$$\left[(\boldsymbol{\sigma} \cdot \mathbf{\Pi})\boldsymbol{\sigma} \cdot \boldsymbol{\nabla} V - (\boldsymbol{\sigma} \cdot \boldsymbol{\nabla} V)(\boldsymbol{\sigma} \cdot \mathbf{\Pi})\right]\varphi = \left[-i\hbar \triangle V + 2i\boldsymbol{\sigma} \cdot (\mathbf{\Pi} \times \boldsymbol{\nabla} V)\right]\varphi\,. \tag{41}$$

By denoting

$$A = (\boldsymbol{\sigma} \cdot \mathbf{\Pi}) \qquad B = (E' - qV)$$
$$C = \boldsymbol{\sigma} \cdot \boldsymbol{\nabla} V \qquad D = -i\hbar \triangle V + 2i\boldsymbol{\sigma} \cdot (\mathbf{\Pi} \times \boldsymbol{\nabla} V)$$

we can obviously write eq.(39) as

$$AB - BA = i\hbar qC$$

and eq.(41) as

$$AC - CA = D\,.$$

After quite long calculations, we obtain the following relation:

$$\begin{aligned}(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2(E' - qV) &+ (E' - qV)(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2 \\ &= q\hbar^2 \triangle V - 2q\hbar\boldsymbol{\sigma} \cdot (\mathbf{\Pi} \times \boldsymbol{\nabla} V) + 2(\boldsymbol{\sigma} \cdot \mathbf{\Pi})(E' - qV)(\boldsymbol{\sigma} \cdot \mathbf{\Pi})\,.\end{aligned} \tag{42}$$

By developing eq.(38) at the second order $\frac{v^2}{c^2}$ and injecting eq.(42) into eq.(38), then developing $(\boldsymbol{\sigma} \cdot \mathbf{\Pi})^2$ at the order $\frac{v^2}{c^2}$, with $q = -e$ the electronic charge and $\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$, we finally obtain the **developed Dirac equation** for an electron (Spin 1/2) interacting with an EM field defined by $\mathbf{A}, V$:

$$\boxed{\left[\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) + \frac{e}{mc}\mathbf{S} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) - \frac{p^4}{8m^3c^2} - \frac{e\hbar}{8m^2c^2} \triangle V - \frac{e}{2m^2c^2}\mathbf{S} \cdot (\boldsymbol{\nabla} V \times \mathbf{p}) - eV\right]\Psi = E'\Psi}\,. \tag{43}$$

6 Conclusion

Starting from "simple" physics (Hamiltonian, Lagrangian of a particle -electron- under the influence of an EM field), we saw the starting point that led Dirac towards writing a new equation. At the end of this report, we developed a **time-independent** eigenvalue equation which is a good starting point for the description of the properties in **atomic** and **molecular physics**. We saw that, after an attempt of O.Klein and W.Gordon to describe spin-0 particles and taking into account the relativistic effects, many improvements have been done leading us to a generalist equation for the behavior of an electron (Spin 1/2) where we notice the presence of the field **B**. Dirac's theory shines by its **consistency with the previous models** and by its "natural" introduction of the spin. When developed with no EM field, Dirac equation reduces to the **time-independent Schrödinger equation**. One of the most important things to notice is that most of the developments in the last part are limited to the second order $\frac{v^2}{c^2}$. Indeed, in the cases we studied here, the terms of higher orders are negligible.

This report is a summary of what I read and calculated during a week of "Supervised research work". The most important thing in the realisation of this report was reading, many books are listed in the references section. None of this work would have been possible (or way harder) without the help of **Pr. Claude Leroy**, who supervised this work. I wish to thank him for the time he allowed me, to talk and clarify some of the hardest parts of this work, giving me a preview of what research in physics looks like. I spent a lot of time in his office discussing about Physics and this week was not limited to the demonstration of the Dirac equation, I also treated theoretical aspects such as sums of angular momenta with the **Clebsch-Gordan** coefficients and the **Wigner-Eckart** theorem. **Pr. Claude Leroy** was very available for me and this whole work conforted me in the idea of becoming a researcher in Physics.

7 Appendix - Useful relations

1.

$$\begin{aligned}(\mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A}))_x &= \dot{y}(\boldsymbol{\nabla} \times \mathbf{A})_z - \dot{z}(\boldsymbol{\nabla} \times \mathbf{A})_y \\ &= \dot{y}(\partial_x A_y - \partial_y A_x) - \dot{z}(\partial_z A_x - \partial_x A_z) \\ &= \dot{y}\partial_x A_y - \dot{y}\partial_y A_x - \dot{z}\partial_z A_x + \dot{z}\partial_x A_z + \dot{x}\partial_x A_x - \dot{x}\partial_x A_x \\ &= \mathbf{v} \cdot \frac{\partial}{\partial x}\mathbf{A} - \mathbf{v} \cdot \boldsymbol{\nabla} A_x \end{aligned} \tag{44}$$

2.

$$[\mathbf{L}, H_{\mathrm{Dirac}}] = [(\mathbf{r} \times \mathbf{p})\mathbb{1}_4, c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2] = i\hbar c(\boldsymbol{\alpha} \times \mathbf{p}) \tag{45}$$

indeed,

$$\begin{aligned}[L_z, H_D] &= c\left[xp_y - yp_x, \sum_\mu \alpha_\mu p_\mu\right] \\ &= i\hbar c(\alpha_x p_y - \alpha_y p_x) \\ &= i\hbar c(\boldsymbol{\alpha} \times \mathbf{p})_z \end{aligned} \tag{46}$$

3. By expliciting $\boldsymbol{\alpha}$ we can demonstrate that

$$(\mathbf{V} \cdot \boldsymbol{\alpha})(\mathbf{W} \cdot \boldsymbol{\alpha}) = \mathbf{V} \cdot \mathbf{W} + i\boldsymbol{\Sigma} \cdot (\mathbf{V} \times \mathbf{W}) \tag{47}$$

4. The demonstration has been done that

$$((\mathbf{p} - qA) \times (\mathbf{p} - qA))_x = iq\hbar(\boldsymbol{\nabla} \times \mathbf{A})_x = iq\hbar B_x \tag{48}$$

Bibliography

[1] Claude Aslangul. *Mécanique quantique : Fondements et Applications*, volume 1. De Boeck Supérieur, p197-210. 2016.

[2] Claude Aslangul. *Mécanique quantique : Fondements et Applications*, volume 2. De Boeck Supérieur, p.968-978. 2015.

[3] Claude Aslangul. *Mécanique quantique : Fondements et Applications*, volume 3. De Boeck Supérieur, p.117-119. 2015.

[4] James H. Smith. *Introduction à la relativité*. Elsevier-Masson, p.124-140. 1997.

[5] Jean-Philippe Pérez. *Relativité : Fondements et Applications*. Dunod, p.85-105. 2016.