The Dirac Equation in the Low Energy Limit

Bachelor of Science Degree in Physics

Björn Linder

Supervisor Michael Bradley

Department of Physics
Umeå University
June 2016

Abstract

The purpose of this work is to give an introduction to relativistic quantum mechanics for the electron. First a review of the Dirac equation and its application to a free electron and an electron in an electromagnetic field will be given. In particular the hydrogen atom is studied. The low energy limit and the lowest order corrections to non-relativistic quantum mechanics are investigated. These corrections include, among others, the gyromagnetic ratio of the electron and spin-orbit coupling in the hydrogen atom. The low energy limit will be studied both in the usual Pauli-Dirac representation and by using the Foldy-Wouthuysen transformation. The Pauli-Dirac representation is used to obtain the correction terms for the energy levels in the hydrogen atom. The Foldy-Wouthuysen transformation is applied to obtain the corresponding Hamiltonian corrections to order α^4.

Sammanfattning

Syftet med det här arbetet är att ge en introduktion till relativistisk kvantmekanik tillämpad på elektronen. Först redogörs Dirac-ekvationen och dess tillämpning hos en fri elektron och en elektron i ett elektromagnetiskt fält. Särskilt studeras ekvationens tillämpning på väteatomen. Den ickerelativistiska gränsen och de lägsta störningstermerna behandlas. Dessa störningstermer inkluderar bland annat det gyromagnetiska kvoten hos elektronen och spin-orbit coupling i väteatomen. Den icke-relativistiska gränsen analyseras både i Pauli-Dirac representationen och genom att använda Foldy-Wouthuysen transformationen. Pauli-Dirac representationen används för att hitta störningstermerna till energinivåerna i väteatomen. Foldy-Wouthuysen transformationen tillämpas för att hitta de motsvarande störningstermerna av ordning α^4 i Hamiltonianen.

Contents

1 Introduction

At the heart of quantum mechanics lies Erwin Schrödinger's famous differential equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \hat{H}\Psi(\boldsymbol{r},t).$$

The solutions obtained by solving the equation successfully predict the behaviour of many quantum mechanical systems. A case which is of great importance and studied in any undergraduate course in quantum mechanics is the case of an electron moving in the Coulumb potential, more specifically the potential produced by a hydrogen nucleus. The solutions for this potential can be obtained exactly but the results fall short of describing the physical world. Experimental data show that there must be more to the story than what is obtained from Schrödinger's equation with a Hamiltonian for an electromagnetic field. The way around this difficulty is to introduce correction terms to the Hamiltonian, corrections that would also adjust the predicted energies. The correction terms of order α^4 (α being the fine structure constant) needed to explain fine structure splittings of the hydrogen spectrum are

$$\begin{aligned}
H_r &= -\frac{1}{8}\frac{(\boldsymbol{p}^2)^2}{m^3c^2}, && \text{(relativistic correction)}\\
H_{so} &= \frac{e}{2m^2c^2}\frac{1}{r}\frac{\partial\Phi}{\partial r}\boldsymbol{S}\cdot\boldsymbol{L}, && \text{(spin-orbit correction)}\\
H_d &= \frac{e\hbar^2}{8m^2c^2}\nabla^2\Phi(\boldsymbol{r}), && \text{(Darwin term)}
\end{aligned}$$

where $\boldsymbol{p}$, $\boldsymbol{S}$ and $\boldsymbol{L}$ are the operators for momentum, spin and the azimuthal quantum number, respectively. $\hbar$ is Planck's reduced constant, e is the elementary charge and Φ is the electric field potential. These correction terms are often introduced with some hand-waving arguments and might not be very easy to motivate. The first correction term can be motivated by the relativistic effects that the electron is subjected to and the second two terms can be motivated by the fact that the electron is a spin-$\frac{1}{2}$ particle. Spin is a property that the Schrödinger equation fail to describe and to do so one would turn to an extension of the equation, the Pauli-Schrödinger equation. The Pauli-Schrödinger equation introduces 2-component solutions and is completely compatible with the spin property of all spin-$\frac{1}{2}$ particles. In the 1920's a search for an equation unifying special relativity and quantum mechanics was taking place. An equation from which, it would turn out, the spin property and all the corrections terms of order α^4 came out naturally.

2 Background

Klein-Gordon Equation

The relativistic analog to Schrödinger's equation would, unlike the Schrödinger equation, meet all the requirements of Einstein's theory of special relativity. An equation today known as the Klein-Gordon equation was proposed as a candidate to by O. Klein, W. Gordon and E. Schrödinger [9, p.115]. Even though the Klein Gordon equation fail to describe the electron, it is a good way to introduce the Dirac Equation.

The energy of a free particle in special relativity is given by the relation

$$E = \pm\sqrt{c^2\boldsymbol{p}^2 + m^2c^4}. \tag{2.1}$$

Searching for a relativistic analog to the Schrödinger equation we could replace E and $\boldsymbol{p}$ with the corresponding operators from quantum mechanics,

$$E \to i\hbar\frac{\partial}{\partial t}, \qquad \boldsymbol{p} \to -i\hbar\nabla. \tag{2.2}$$

Applying this energy operator on a function ψ gives a differential equation,

$$i\hbar\frac{\partial\psi}{\partial t} = \pm\sqrt{-\hbar^2c^2\nabla^2 + m^2c^4}\psi. \tag{2.3}$$

However, the Taylor expansion of the spacial derivative give derivatives to all orders up to infinity. The derivatives of space and time are hence of different orders and they cannot be symmetric, which would be required in special relativity [9, p.118].

If one instead takes the square of equation (2.1) and then substitute the quantum mechanical operators of (2.2) we get

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = -\hbar^2c^2\nabla^2\psi + m^2c^4\psi. \tag{2.4}$$

This is known as the Klein-Gordon equation for a free particle and it can take on a slightly more compact form by introducing the d'Alembertian ($\Box \equiv \frac{1}{c^2}(\partial^2/\partial t^2) - \nabla^2$),

$$-\Box\psi = \frac{m^2c^2}{\hbar^2}\psi. \tag{2.5}$$

The solutions to the Klein Gordon equation for a free particle are the plane wave solutions

$$\psi = Ae^{i(Et-\boldsymbol{p}\cdot\boldsymbol{x})/\hbar}, \tag{2.6}$$

where the energy E can have both positive and negative values as described by eq. (2.1) and hence allows for two separate solutions [9, p.120]. The theory does, however, not include spin which and hence cannot fully describe the electron. However, the major problem with the Klein-Gordon equation appears when studying the analog to the conservation equation of non-relativistic quantum mechanics [8, p.468]. Gathering the terms in eq. (2.5) and multiplying with the complex conjugate wave function ψ^* gives

$$\psi^* \Box \psi + \psi^* \frac{m^2 c^2}{\hbar^2} \psi = 0. \tag{2.7}$$

By subtracting eq. (2.7) by its complex conjugate we get

$$\psi^* \Box \psi - \psi \Box \psi^* = 0, \tag{2.8}$$

which is equivalent to

$$\frac{\partial}{\partial t}\left(\frac{1}{c^2}\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right)\right) + \nabla \cdot \left(\psi \nabla \psi^* - \psi^* \nabla \psi\right) = 0. \tag{2.9}$$

This is of the same form as the continuity equation of non-relativistic quantum mechanics,

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

If multiplied with a factor $\hbar/2mi$ the second term of eq. (2.9) is identical to the current density $\boldsymbol{j}$ in the classical case but the analog to ρ, which classically can be interpreted as the position probability, is not the same. Multiplying with the missing factor we get

$$P(\boldsymbol{x}, t) = \frac{i\hbar}{2mc^2}\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right), \tag{2.10}$$

which allows for negative probabilities. This is a fundamental error and comes from the fact that the Klein-Gordon is a second order differential equation, with respect to time, in which the initial conditions for ψ and $\partial\psi/\partial t$ may be chosen independently [9, p.120]. It was mainly because of this feature the Klein-Gordon equation was rejected.

3 Dirac Equation

Deriving the Dirac Equation

Dirac wanted to find a first order equation describing the problem, a relativistic analog in which the wave functions would still satisfy the energy-momentum relation $E^2 = p^2c^2 + m^2c^4$ of the Klein-Gordon equation. In Dirac's exact words "The general interpretation of non-relativity quantum mechanics is based on the transformation theory, and is made possible by the wave equation of the form $(H - E)\psi = 0$, i.e. being linear in E or $\partial/\partial t$, so that the wave function at any time determines the wave function at any later time" [4, p.612].

Using the definition of the d'Alembertian and the short-hand notation $\frac{\partial}{\partial t} = \partial_t$ equation (2.5) can be written as

$$\left(\nabla^2 - \frac{1}{c^2}\partial_t^2\right)\psi = \frac{m^2c^2}{\hbar^2}\psi. \tag{3.1}$$

One way to solve the problem of the second order derivatives would be to factorise the d'Alembertian operator into a square of some operator and apply this new operator once to the wave function. Introducing the notation $\frac{\partial}{\partial x} = \partial_x$, $\frac{\partial}{\partial y} = \partial_y$ and $\frac{\partial}{\partial t} = \partial_t$, the factorisation

$$\left(\nabla^2 - \frac{1}{c^2}\partial_t^2\right) = \left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)^2, \tag{3.2}$$

can work if the right requirements on the coefficients A, B, C, D are chosen. They have to be anti-commutative so that the cross derivatives disappear and they also need to be their own inverse to preserve the second order derivatives. In other words we need that

$$AB + BA = 0, \quad BC + CB = 0, \quad \text{etc.} \qquad \text{and} \qquad A^2 = B^2 = C^2 = D^2 = 1. \tag{3.3}$$

The requirements can be met if the coefficients are matrices. In fact these matrices are closely related to the Dirac γ-matrices that Dirac defined in his slightly different derivation of the equation,

$$A = i\gamma^1, B = i\gamma^2, C = i\gamma^3, D = \gamma^0. \tag{3.4}$$

The gamma matrices are 4×4 hermitian matrices built up of the Pauli matrices[1] of non-relativistic quantum mechanics,

$$\gamma^k = \begin{bmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{bmatrix} \quad \text{and} \quad \gamma^0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \tag{3.5}$$

[1] Defined as $\left(\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.\right)$

where I is the 2×2 identity matrix. The 0's in the matrices are to be interpreted as the 2×2 zero matrices. Some properties of the Dirac matrices are

$$[\gamma^\nu, \gamma^\mu]_+ = 2g^{\mu\nu} I, \tag{3.6a}$$

$$(\gamma^0)^2 = I, \quad (\gamma^k)^2 = -I, \tag{3.6b}$$

$$(\gamma^0)^\dagger = \gamma^0, \quad (\gamma^k)^\dagger = -\gamma^k, \tag{3.6c}$$

where $[\gamma^\nu, \gamma^\mu]_+$ denotes the anti-commutator $\gamma^\nu\gamma^\mu + \gamma^\mu\gamma^\nu$ and $g^{\mu\nu}$ is the metric from special relativity. Introducing Einstein's sum convention and the space-time notation from special relativity, $x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$, we can rewrite the factorised operator.[2] We get that

$$\left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)^2 = \left(i\gamma^\mu \frac{\partial}{\partial x^\mu}\right)^2. \tag{3.7}$$

Applying this factor once to a wave function should in accordance with (3.1) give the eigenvalue equation

$$\left(i\gamma^\mu \frac{\partial}{\partial x^\mu}\right)\psi = \frac{mc}{\hbar}\psi, \tag{3.8}$$

which is the Dirac Equation for a free particle in its manifest covariant form.[3] There is another common way of expressing the Dirac Equation and to do so we notice that the gamma matrices are related by a common matrix factor $\beta = \gamma^0$. Hence they can be written as

$$\gamma^1 = \beta\alpha_1, \quad \gamma^2 = \beta\alpha_2, \quad \gamma^3 = \beta\alpha_3, \quad \gamma^0 = \beta. \tag{3.9}$$

Substituting (3.9) into the Dirac Equation (3.8) we get

$$\left(i\beta\alpha_1\partial_x + i\beta\alpha_2\partial_y + i\beta\alpha_3\partial_z + \frac{i}{c}\beta\partial_t\right)\psi = \frac{mc}{\hbar}\psi. \tag{3.10}$$

Multiplying with $c\hbar\beta$ on the left and rearranging the terms leaves

$$-\hbar c\left(i\alpha_1\partial_x + i\alpha_2\partial_y + i\alpha_3\partial_z - \frac{mc}{\hbar}\beta\right)\psi = i\hbar\partial_t\psi, \tag{3.11}$$

where we have used that β is an invertible matrix with $\beta^{-1} = \beta$. We can simplify further to get

$$(c\boldsymbol{\alpha}\cdot(-i\hbar\nabla) + mc^2\beta)\psi = i\hbar\partial_t\psi \tag{3.12}$$

or equivalently

$$(c\boldsymbol{\alpha}\cdot\boldsymbol{p} + mc^2\beta)\psi = E\psi. \tag{3.13}$$

[2] See Appendix A for a short introduction to the metric and the notation of special relativity.

[3] There are different ways of representing the gamma matrices, which causes the Dirac Equation to take on different forms. This representation is based on notation by Bjorken, Drell and Schweber.

This equation is of the same form as the Schrödinger equation, $\hat{H}\psi = E\psi$, with the Hamiltonian given by

$$\hat{H} = c\boldsymbol{\alpha} \cdot \boldsymbol{p} + mc^2\beta. \tag{3.14}$$

The hermitian matrices β and α are

$$\beta = \gamma^0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \qquad \alpha_k = \begin{bmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{bmatrix} = \gamma^0\gamma^k \qquad (k = 1, 2, 3), \tag{3.15}$$

and satisfy

$$\begin{gathered} \alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = 1, \\ [\alpha_k, \beta]_+ = 0, \qquad \beta^2 = 1, \qquad [\alpha_k, \alpha_l]_+ = 2\delta_{kl}, \end{gathered} \tag{3.16}$$

where $[\alpha_i, \alpha_j]_+$ is the notation for the anticommutator $\alpha_i\alpha_j + \alpha_j\alpha_i$ and δ_{kl} is the Kronecker delta function.

Conservation Equation

To obtain a conservation equation we first multiply (3.12) on the left with the conjugate transpose $\psi^\dagger$

$$\psi^\dagger\Big(i\hbar\partial_t\psi + c\boldsymbol{\alpha} \cdot (i\hbar\nabla\psi) - mc^2\beta\psi\Big) = 0. \tag{3.17}$$

Introducing the hermitian adjoint version of (3.12) and multiplying it on the right with ψ we get

$$\Big(-i\hbar\partial_t\psi^\dagger - c\boldsymbol{\alpha} \cdot (i\hbar\nabla\psi^\dagger) - mc^2\beta\psi^\dagger\Big)\psi = 0, \tag{3.18}$$

where we have used that the matrices α and β are hermitian. Taking the difference of them and setting it equal to zero gives us the equation

$$\bigg(\psi^\dagger\partial_t\psi + (\partial_t\psi^\dagger)\psi\bigg) + c\bigg(\psi^\dagger\alpha_k\partial_{x^k}\psi + \alpha_k(\partial_{x^k}\psi^\dagger)\psi\bigg) = 0, \tag{3.19}$$

which can be interpreted as a conservation equation if $\rho = \psi^\dagger\psi$ and the current density $j_k = \psi^\dagger\alpha_k\partial_k\psi$ [8, p.477].

4 Solutions for the Free Particle

Looking at the Hamiltonian $\hat{H} = c\boldsymbol{\alpha} \cdot \boldsymbol{p} + mc^2\beta$, which is a 4×4 matrix, it must be so that any wave-function ψ satisfying the Dirac equation must be a four element row vector,

$$\Psi(\boldsymbol{x}, t) = \begin{bmatrix} \psi_1(\boldsymbol{x}, t) \\ \psi_2(\boldsymbol{x}, t) \\ \psi_3(\boldsymbol{x}, t) \\ \psi_4(\boldsymbol{x}, t) \end{bmatrix}. \tag{4.1}$$

This 4-element vector solution is called a bispinnor or a Dirac spinor. Solutions for a particle not subjected to any potential can be found by making the ansatz

$$\Psi(\boldsymbol{x}, t) = u(\boldsymbol{p}) e^{i[(\boldsymbol{p} \cdot \boldsymbol{x}/\hbar) - (Et/\hbar)]}, \tag{4.2}$$

where $u(\boldsymbol{p})$ is a 4×1 column vector.

Stationary Particle

The most simple solution to the Dirac equation is found in the case of a free, stationary particle with $p = 0$. The Hamiltonian then reduces to $\hat{H} = mc^2\beta$ and the equation becomes

$$i\hbar \frac{\partial \psi}{\partial t} = mc^2 \beta \psi. \tag{4.3}$$

The plane-wave ansatz reduces down to $u_j(\mathbf{0}) e^{-imc^2 t/\hbar}$ (if $E = +mc^2$) and plugging it into the Dirac equation we get the system of equations

$$\frac{mc}{\hbar} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} u_A \\ u_B \end{bmatrix} = \frac{mc}{\hbar} \begin{bmatrix} u_A \\ u_B \end{bmatrix}, \tag{4.4}$$

where the u_j's are split into the two 2×1 column vectors u_A and u_B. This requires that u_B (u_3 and u_4) is zero. However, if instead the negative energy was chosen u_B (u_1 and u_2) has to vanish. For either of these cases the solutions for u_A and u_B are the Pauli spinors

$$\chi_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{or} \quad \chi_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4.5}$$

Hence there are four solutions, with two corresponding to positive energy and two corresponding to negative energy [7, p.89-90]. The solutions are

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} e^{-imc^2 t/\hbar}, \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} e^{-imc^2 t/\hbar}, \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} e^{imc^2 t/\hbar}, \qquad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} e^{imc^2 t/\hbar}. \tag{4.6}$$

Particle in Motion

We could now start looking for solutions for the free particle with $p > 0$. The following derivation is based on a derivation made by Franz Schwabl in his book *Advanced Quantum Mechanics* [9, p.147-149]. We start by introducing the 4-position vector x and 4-momentum vector k.

$$x \equiv (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \boldsymbol{x}) \tag{4.7a}$$

$$k \equiv (k^0, k^1, k^2, k^3) = (E/c, p_x, p_y, p_z) = (E/c, \boldsymbol{k}) \tag{4.7b}$$

Assuming that the solutions will be in accordance with the stationary particle we will try to find independent solutions corresponding to the positive and negative energies. The four 4-component solutions are assumed to be in the form

$$\Psi^{(+)} = u_r^{(+)}(k) e^{-ik\cdot x/\hbar}, \quad \Psi^{(-)} = u_r^{(-)}(k) e^{ik\cdot x/\hbar}, \qquad (r = 1, 2). \tag{4.8}$$

We will also introduce the Feynman slash notation

$$\not{\nu} \equiv \gamma \cdot \nu = \gamma^\mu \nu_\mu = \gamma_0 \nu_0 - \boldsymbol{\gamma} \cdot \boldsymbol{\nu}. \tag{4.9}$$

The Dirac Equation eq. (3.8) can then be written in the compact form

$$\left(i\not{\partial} - \frac{mc}{\hbar} \right) \psi = 0. \tag{4.10}$$

Plugging $\psi_r^{(+)}$ and $\psi_r^{(-)}$ into the Dirac Equation (4.10) we get the two equations

$$(\not{k} - mc) u_r^{(+)}(k) = 0 \qquad (\not{k} + mc) u_r^{(-)}(k) = 0. \tag{4.11}$$

The two factors are each others conjugate,

$$(\not{k} - mc)(\not{k} + mc) = \not{k}\not{k} - (mc)^2, \tag{4.12}$$

and if $\not{k}\not{k}$ is investigated more closely we find that

$$\not{k}\not{k} = k_\mu \gamma^\mu k_\nu \gamma^\nu = k_\mu k_\nu \frac{1}{2} [\gamma^\mu, \gamma^\nu]_+ = k_\mu k_\nu g^{\mu\nu}. \tag{4.13}$$

This is just the invariant inner product of the 4-momentum,

$$k_\mu k^\mu = \frac{E^2}{c^2} - \boldsymbol{k}^2 = m^2 c^2. \tag{4.14}$$

Substituting (4.14) into eq. (4.12) we see that it is equal to zero. Hence equations (4.11) are fulfilled if the coefficients $u_r^{(+)}$ and $u_r^{(-)}$ contain the corresponding conjugate factor. The solutions are simply the coefficients from the free particle at rest times these factors

$$u_r^{(+)}(k) = N(\not{k} + mc) u_r(\mathbf{0}) \qquad u_r^{(-)}(k) = N(\not{k} - mc) u_{r+2}(\mathbf{0}) \tag{4.15}$$

Once again introducing the Pauli spinors we can write the equations for the coefficients as

$$u_r(k) = N(\not{k} + mc)\begin{bmatrix}\chi_r\\0\end{bmatrix} \quad \text{and} \quad v_r(k) = N(\not{k} - mc)\begin{bmatrix}0\\\chi_r\end{bmatrix}, \tag{4.16}$$

where the 0's are abbreviations for the 2×1 zero matrices. The calculations are straightforward,

$$\begin{aligned}
\not{k}\begin{bmatrix}\chi_r\\0\end{bmatrix} &= \left(k^0\begin{bmatrix}I & 0\\0 & -I\end{bmatrix} - k^i\begin{bmatrix}0 & \sigma_i\\-\sigma_i & 0\end{bmatrix}\right)\begin{bmatrix}\chi_r\\0\end{bmatrix} = \\
&= \begin{bmatrix}k^0 I & -k^i\sigma_i\\k^i\sigma_i & k^0 I\end{bmatrix}\begin{bmatrix}\chi_r\\0\end{bmatrix} = \begin{bmatrix}\frac{E}{c}\chi_r\\\boldsymbol{\sigma}\cdot\boldsymbol{k}\chi_r\end{bmatrix}\\
-\not{k}\begin{bmatrix}0\\\chi_r\end{bmatrix} &= \left(k^0\begin{bmatrix}-I & 0\\0 & +I\end{bmatrix} + k^i\begin{bmatrix}0 & \sigma_i\\-\sigma_i & 0\end{bmatrix}\right)\begin{bmatrix}0\\\chi_r\end{bmatrix} = \\
&= \begin{bmatrix}-k^0 I & k^i\sigma_i\\-k^i\sigma_i & k^0 I\end{bmatrix}\begin{bmatrix}0\\\chi_r\end{bmatrix} = \begin{bmatrix}\boldsymbol{k}\cdot\boldsymbol{\sigma}\chi_r\\\frac{E}{c}\chi_r\end{bmatrix}.
\end{aligned}$$

This give the coefficients

$$u_r^{(+)}(k) = N\begin{bmatrix}(\frac{E}{c} + mc)\chi_r\\\boldsymbol{\sigma}\cdot\boldsymbol{k}\chi_r\end{bmatrix} \quad \text{and} \quad u_r^{(-)}(k) = N\begin{bmatrix}\boldsymbol{k}\cdot\boldsymbol{\sigma}\chi_r\\(\frac{E}{c} - mc)\chi_r\end{bmatrix}. \tag{4.17}$$

Using the fact that

$$\boldsymbol{\sigma}\cdot\boldsymbol{k} = \begin{bmatrix}p_z & p_x - ip_y\\px + ip_y & -p_z\end{bmatrix} \tag{4.18}$$

we get the solutions

$$u_1^{(+)} = N\begin{bmatrix}(E + mc^2)/c\\0\\p_z\\(p_x + ip_y)\end{bmatrix}, \quad u_2^{(+)} = N\begin{bmatrix}0\\(E + mc^2)/c\\(p_x - ip_y)\\-p_z\end{bmatrix}, \tag{4.19a}$$

$$u_1^{(-)} = N\begin{bmatrix}p_z\\(p_x + ip_y)\\(E - mc^2)/c\\0\end{bmatrix}, \quad u^{(-)} = N\begin{bmatrix}-(p_x - ip_y)\\-p_z\\0\\(E - mc^2)/c\end{bmatrix}. \tag{4.19b}$$

The solutions are more commonly written as

$$u_1 = N' \begin{bmatrix} 1 \\ 0 \\ cp_z/(E+mc^2) \\ c(p_x+ip_y)/(E+mc^2) \end{bmatrix}, \qquad u_2 = N' \begin{bmatrix} 0 \\ 1 \\ c(p_x-ip_y)/(E+mc^2) \\ -cp_z/(E+mc^2) \end{bmatrix}, \tag{4.20a}$$

$$v_1 = N' \begin{bmatrix} -cp_z/(|E|+mc^2) \\ -c(p_x+ip_y)/(|E|+mc^2) \\ 1 \\ 0 \end{bmatrix}, \qquad v_2 = N' \begin{bmatrix} -(p_x-ip_y)/(|E|+mc^2) \\ cp_z/(|E|+mc^2) \\ 0 \\ 1 \end{bmatrix}. \tag{4.20b}$$

We have used that the energy in the $u_r^{(-)}$ solutions have a negative sign relative to the energies in $u_r^{(+)}$. It is straightforward to show that the $u^{(-)}$'s and $u^{(+)}$'s are orthogonal to each other for a given $\boldsymbol{p}$ as well as

$$u_r^\dagger(\boldsymbol{p})u_{r'}(\boldsymbol{p}) = 0 \quad \text{for } r \neq r' \tag{4.21}$$

When it comes to the normalisation there are two different approaches [7, p.93]. The first is that

$$u_r^\dagger(\boldsymbol{p})u_r(\boldsymbol{p}) = 1 \tag{4.22}$$

which implies that

$$N' = \sqrt{|E| + mc^2)/2|E|}. \tag{4.23}$$

The second convention is that

$$u_r^\dagger(\boldsymbol{p})u_r(\boldsymbol{p}) = |E|/mc^2, \tag{4.24}$$

which gives

$$N' = \sqrt{(|E| + mc^2)/2mc^2}. \tag{4.25}$$

5 Spin and Spin Orientation

For either sign of the energy there are two different solutions. We introduce the 4×4 spin operator

$$\boldsymbol{S} \equiv \frac{\hbar}{2}\boldsymbol{\Sigma} \equiv \frac{\hbar}{2}\begin{bmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{bmatrix}, \tag{5.1}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli spin matrices. In this sense it is really the 4×4 extension of the classical quantum mechanic spin operator $\hat{\boldsymbol{S}}_c = (\hbar/2)\boldsymbol{\sigma}$. For a particle with $\boldsymbol{p} = 0$, which was the first of the two cases we solved in the previous section, the different solutions correspond to different spin states, as they are eigenspinors to the operator $\hat{S}_3$. Applying the operator to these solutions we get the eigenvalues

$$\hat{S}_3 u_1 = (+1)\frac{\hbar}{2}u_1, \quad \hat{S}_3 u_2 = (-1)\frac{\hbar}{2}u_2, \tag{5.2}$$

$$\hat{S}_3 v_1 = (+1)\frac{\hbar}{2}v_1, \quad \hat{S}_3 v_2 = (-1)\frac{\hbar}{2}v_2. \tag{5.3}$$

The first solutions for both signs of the energy hence correspond to spin +1 and the second solutions of each set correspond to spin −1. For the free particle with $\boldsymbol{p} \neq 0$ this is not true and applying the operator $\hat{S}_3$ to the solutions will not give an equivalent eigenvalue [7, p.92-93]. It will be possible if we assume that the momentum is parallel to the spin, which can be seen as following. If the momentum is along the z-axis we get

$$\hat{S}_3 u_1(\boldsymbol{p} = p_z) = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} N' \begin{bmatrix} 1 \\ 0 \\ cp_z/(E + mc^2) \\ 0 \end{bmatrix} = (+1)(\frac{\hbar}{2})u_1,$$

$$\hat{S}_3 u_2(\boldsymbol{p} = p_z) = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} N' \begin{bmatrix} 0 \\ 1 \\ 0 \\ -cp_z/(E_+ + mc^2) \end{bmatrix} = (-1)(\frac{\hbar}{2})u_2.$$

Likewise the eigenvalues are $\hbar/2$ and $-\hbar/2$ for v_1 and v_2 respectively. Hence we can define the helicity operator

$$\Lambda \equiv \boldsymbol{\Sigma} \cdot \frac{\boldsymbol{p}}{|\boldsymbol{p}|}, \tag{5.4}$$

of which the Dirac spinor for a free particle in motion is an eigenfunction [7, p. 93]. In other words the spinors are eigenspinors to the spin operator in the direction of the momentum. The states corresponding to (+1) eigenvalues as right-handed states and the states corresponding to (−1) eigenvalues as left-handed states. The right-handed states are those with motion and spin in the same direction and the left-handed states are those with motion and spin in opposite directions.

It can be shown that neither of the operators $\boldsymbol{L} \equiv \boldsymbol{x} \times \boldsymbol{p}$ and $\boldsymbol{S} \equiv \frac{\hbar}{2}\boldsymbol{\Sigma}$ commute with the Hamiltonian so that in Heisenberg's picture they are not constants of motion and hence not conserved.[4] The total angular momentum $\boldsymbol{J} = \boldsymbol{x} \times \boldsymbol{p} + \frac{\hbar}{2}\boldsymbol{\Sigma}$ is however a constant of motion in the presence of a spherically symmetrical potential and in the case of no potential [7, p.122]. In classical mechanics we can specify the orientation of the spin relative to the total angular momentum with the operator $\boldsymbol{\sigma} \cdot \boldsymbol{J}$. In solving the Dirac equation for a hydrogen potential it will be of interest to have a relativistic analog to such an operator. We introduce the operator $\beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}$ which is quite similar to the non-relativistic operator. Its commutation relation with the free Hamiltonian (equation 3.14) is

$$[H, \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}] = \frac{\hbar}{2}(-2c\beta\boldsymbol{\alpha} \cdot \boldsymbol{p}) \tag{5.5}$$

[7, p.122]. We also have the following commutation relation, which is easier to verify,

$$[H, \beta] = c\boldsymbol{\alpha} \cdot \boldsymbol{p}\beta - \beta c\boldsymbol{\alpha} \cdot \boldsymbol{p} = -2c\beta\boldsymbol{\alpha} \cdot \boldsymbol{p}. \tag{5.6}$$

Based on these two equation a new operator K that commutes with H can be introduced,

$$K \equiv \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J} - \frac{\hbar}{2}\beta = \beta(\boldsymbol{\Sigma} \cdot \boldsymbol{L} + \hbar). \tag{5.7}$$

K does not only commute with H but it also commutes with $\boldsymbol{J}$, since it commutes with the two matrix terms of $\boldsymbol{J}$, β and $\boldsymbol{\Sigma} \cdot \boldsymbol{L}$. Because of this we can construct functions that are eigenfunctions of $H, K, \boldsymbol{J}^2$ and J_z. Further important relations between the eigenvalues k and $j(j+1)$ can be derived. We have that

$$\begin{aligned} K^2 = \beta(\boldsymbol{\Sigma} \cdot \boldsymbol{L} + \hbar)\beta(\boldsymbol{\Sigma} \cdot \boldsymbol{L} + \hbar) &= (\boldsymbol{\Sigma} \cdot \boldsymbol{L} + \hbar)^2 \\ &= (\boldsymbol{\Sigma} \cdot L)^2 + 2(\boldsymbol{\Sigma} \cdot \boldsymbol{L})\hbar + \hbar^2. \end{aligned} \tag{5.8}$$

To progress from here we will introduce the fifth gamma-matrix

$$\gamma^5 \equiv \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{5.9a}$$

which anti-commutes with the other gamma-matrices

$$[\gamma^5, \gamma^\mu]_+ = 0. \tag{5.10a}$$

[4] See Appendix B for more on Heisenberg's picture and the derivation of this.

The components of $\mathbf{\Sigma}$ can then be expressed as (inspired by [2], *Solution of the Dirac Equation for Hydrogen*)

$$\begin{aligned}\Sigma_1 &= \gamma^5\alpha_1 = \gamma^5(\gamma^0\gamma^1) = -i(\gamma^2\gamma^3),\\ \Sigma_2 &= \gamma^5\alpha_2 = \gamma^5(\gamma^0\gamma^2) = -i(\gamma^3\gamma^1),\\ \Sigma_3 &= \gamma^5\alpha_3 = \gamma^5(\gamma^0\gamma^3) = -i(\gamma^1\gamma^2).\end{aligned} \tag{5.11}$$

This is verified by using the associative property of matrix multiplication, taking into consideration that the gamma-matrices are anti-commutative and 'flipping two over' at a cost of changing sign. Because of (5.11) we can write

$$\Sigma_i\Sigma_i = 1, \tag{5.12a}$$

$$\Sigma_i\Sigma_j = i\epsilon_{ijk}\Sigma_k. \tag{5.12b}$$

where ϵ is the Levi-Civita symbol.[5] Using this and that $\boldsymbol{L}\times\boldsymbol{L} = i\hbar\boldsymbol{L}$ [7, p.123] we can rewrite the first term of eq. (5.8) as

$$\begin{aligned}(\mathbf{\Sigma}\cdot L)^2 &= \Sigma_i L_i \Sigma_j L_j(\delta_{ij} + i\epsilon_{ijk}\Sigma_k)\\ &= \boldsymbol{L}^2 + i\mathbf{\Sigma}\cdot(\boldsymbol{L}\times\boldsymbol{L})\\ &= \boldsymbol{L}^2 - \hbar\mathbf{\Sigma}\cdot\boldsymbol{L}.\end{aligned}$$

The operator can now be written as

$$K^2 = \boldsymbol{L}^2 + \hbar\mathbf{\Sigma}\cdot\boldsymbol{L} + \hbar^2. \tag{5.14}$$

We will now make use of that the equation above is very similar to another operator, namely $\boldsymbol{J}^2$.

$$\begin{aligned}\boldsymbol{J}^2 &= \boldsymbol{L}^2 + \hbar\mathbf{\Sigma}\cdot\boldsymbol{L} + 3\hbar/4\\ &= K^2 - \hbar^2/4\end{aligned} \tag{5.15}$$

and in other words the eigenvalues must be connected as

$$k^2\hbar^2 = j(j+1)\hbar^2 + \frac{\hbar^2}{4} = \big(j + \frac{1}{2}\big)^2\hbar^2. \tag{5.16}$$

The eigenvalues of the K operator must therefore be nonzero integers

$$k = \pm\big(j + \frac{1}{2}\big). \tag{5.17}$$

The sign of these eigenvalues determine whether the spin of the particle is parallel or anti-parallel to the angular momentum. A negative sign implies an anti-parallel state and a positive sign a parallel state [7, p.123].

[5] The Levi-Civita symbol takes on the values $+1, -1$ or 0. It will be $+1$ if the sub-indexes are in an cyclic order, for example 123 or 312, and -1 for the odd permutations of 123. Levi-Civita therefor has the anti-symmetric property that interchanging two indexes changes it's sign. If any two indexes are the same it takes on the value zero.

6 Solutions for the Hydrogen Atom

The following derivation is based mainly on a derivation made by J.J. Sakurai in his book *Advanced Quantum Mechanics* [7, p.122-131]. The Hydrogen atom is the only atomic potential for which the Dirac equation can be solved exactly.

In the presence of a spherically symmetrical electromagnetic potential (with no magnetic vector potential, $\boldsymbol{A} = 0$) the Hamiltonian takes the form

$$H = c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2 + V(r) \tag{6.1}$$

where the potential that an electron is subject to in a hydrogen atom is well known to be

$$V(r) = \frac{-e^2}{4\pi\epsilon_0 r}.$$

Obtaining the Radial Equations

Splitting the Dirac spinor Ψ into two 2×1 spinors ψ_A and ψ_B and applying the operator

$$K = \beta(\boldsymbol{\Sigma}\cdot\boldsymbol{L} + \hbar) = \begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{L} + \hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\boldsymbol{L} + \hbar \end{bmatrix} \tag{6.2}$$

introduced in the previous section gives two equations

$$(\boldsymbol{\sigma}\cdot\boldsymbol{L} + \hbar)\psi_A = -k\hbar\psi_A, \qquad (\boldsymbol{\sigma}\cdot\boldsymbol{L} + \hbar)\psi_B = k\hbar\psi_B. \tag{6.3}$$

At the same time, applying the operator $\boldsymbol{J}^2$ of which ψ is also an eigenfunction gives the equations

$$\boldsymbol{J}^2\psi_{A,B} = (\boldsymbol{L} + \hbar\boldsymbol{\sigma}/2)^2\psi_{A,B} = j(j+1)\hbar^2\psi_{A,B} \quad , \quad j = \tfrac{1}{2}, \tfrac{3}{2}, \ldots \tag{6.4}$$

From this we can conclude that while ψ is not an eigenfunction of $\boldsymbol{L}^2$, the components ψ_A and ψ_B can be, as can be seen by expanding the middle term in (6.4). Using the unitary property of the Pauli matrices, $\sigma_i\sigma_i = I$, we can through (6.4) also obtain that when $\boldsymbol{L}^2$ acts on a two-component state vector it is equivalent to $\boldsymbol{J}^2 - \hbar\boldsymbol{\sigma}\cdot\boldsymbol{L} - 3\hbar^2/4$. Applying the operator $\boldsymbol{L}^2$ to $\psi_{A,B}$ must therefore give the relations

$$l_A(l_A+1) = j(j+1) + k + \frac{1}{4} \quad \text{and} \quad l_B(l_B+1) = j(j+1) - k + \frac{1}{4}, \tag{6.5}$$

if the eigenvalues of $\boldsymbol{L}^2$, $l(l+1)$, are assumed to have a factor $\hbar^2$. From equations (6.5) and (5.17) it is possible to determine the two possible values of l_A and l_B for a given k.

$$\begin{aligned} k = j + \frac{1}{2} \quad &\Longrightarrow \quad l_A = j + \frac{1}{2}, \qquad l_B = j - \frac{1}{2}. \\ k = -(j + \frac{1}{2}) \quad &\Longrightarrow \quad l_A = j - \frac{1}{2}, \qquad l_B = j + \frac{1}{2}. \end{aligned} \tag{6.6}$$

In other words the value of l_A and l_B for a given j differ with a value of 1. We make an ansatz for the components of the Dirac spinor

$$\Psi = \begin{bmatrix} \psi_A \\ \psi_B \end{bmatrix} = \begin{bmatrix} g(r) Y_{jl_A}^{m_j} \\ if(r) Y_{jl_B}^{m_j}, \end{bmatrix} \tag{6.7}$$

where the i is there to make the solution real and the functions Y are linear combinations of the spherical harmonics

$$Y_{jl}^{m_j} = \sqrt{\frac{l + m_j + \frac{1}{2}}{2l+1}} Y_l^{m_j - 1/2} \chi_1 + \sqrt{\frac{l - m_j + \frac{1}{2}}{2l+1}} Y_l^{m_j + 1/2} \chi_2 \qquad \left(\text{for } j = l + \tfrac{1}{2}\right), \tag{6.8}$$

$$Y_{jl}^{m_j} = \sqrt{\frac{l - m_j + \frac{1}{2}}{2l+1}} Y_l^{m_j - 1/2} \chi_1 + \sqrt{\frac{l + m_j + \frac{1}{2}}{2l+1}} Y_l^{m_j + 1/2} \chi_2 \qquad \left(\text{for } j = l - \tfrac{1}{2}\right), \tag{6.9}$$

where χ_1 and χ_2 are the Pauli spinors as defined in (4.5). The superindex m_j is the eigenvalue of the operator J_z calculated from

$$J_z \psi_{A,B} = (L_3 + \frac{\hbar}{2}\sigma_3)\psi_{A,B} = m_j \hbar \psi_{A,B} \quad , \quad m_j = -j, \ldots, j. \tag{6.10}$$

The functions $Y_{jl}^{m_j}$ are normalized eigenfunctions of $\boldsymbol{J}^2, \boldsymbol{L}^2, J_z$ and $\boldsymbol{S}^2$ as required. To continue we look at the parity of the functions we have introduced. For a given k the orbital parities of $Y_{jl_A}^{m_j}$ and $Y_{jl_B}^{m_j}$ are opposite. The spherical harmonics transform under a parity inversion as [9, p.170]

$$Y_{jl}^{m_j}(-\boldsymbol{x}) = (-1)^l Y_{jl}^{m_j}(\boldsymbol{x}) \tag{6.11}$$

and hence, looking at (6.6), we can see that ψ_A and ψ_B correspond to opposite parities. The operator $\frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r}$ flips the parity of the spherical harmonics and hence it changes l. It turns out that

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r} Y_{jl_A}^{m_j} = -Y_{jl_B}^{m_j}. \tag{6.12}$$

The square of the operator (i.e. applying the operator twice) reduce down to 1. In other words the operator will change l between l_A and l_B, which will prove useful.

Writing the Hamiltonian (6.1) in its matrix form [9, p.172]

$$\hat{H} = \begin{bmatrix} mc^2 - V(r) & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -mc^2 - V(r) \end{bmatrix}, \tag{6.13}$$

and substituting the Dirac spinor ψ into the Dirac equation we get the equations

$$c(\boldsymbol{\sigma}\cdot\boldsymbol{p})\psi_B = (E - V(r) - mc^2)\psi_A \tag{6.14a}$$
$$c(\boldsymbol{\sigma}\cdot\boldsymbol{p})\psi_A = (E - V(r) + mc^2)\psi_B \tag{6.14b}$$

The parity operator $\frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r}$ comes in by rewriting the scalar product

$$\boldsymbol{\sigma}\cdot\boldsymbol{p} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r^2}(\boldsymbol{\sigma}\cdot\boldsymbol{x})(\boldsymbol{\sigma}\cdot\boldsymbol{p}) = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r^2}\left(- i\hbar r\frac{\partial}{\partial r} + i\boldsymbol{\sigma}\cdot\boldsymbol{L}\right). \tag{6.15}$$

Substituting this as well as the 2×1 component vectors for ψ_A and ψ_B we can calculate the left-hand sides of equations (6.14),

$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})\psi_B = -\hbar\frac{df}{dr}Y^{m_j}_{jl_A} - \frac{(1-k)\hbar}{r}fY^{m_j}_{jl_A}, \tag{6.16a}$$
$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})\psi_A = i\hbar\frac{dg}{dr}Y^{m_j}_{jl_B} + i\frac{(1+k)\hbar}{r}gY^{m_j}_{jl_B}. \tag{6.16b}$$

If substituted back into (6.14) the combinations of spherical harmonics can be removed, leaving just

$$c\hbar\frac{df}{dr} + \frac{(1-k)\hbar c}{r}f = -(E - V - mc^2)g, \tag{6.17a}$$
$$c\hbar\frac{dg}{dr} + \frac{(1+k)\hbar c}{r}g = (E - V + mc^2)f. \tag{6.17b}$$

For convenience we introduce $F(r) = rf(r)$ and $G(r) = rg(r)$ and arrive at the radial equations

$$\hbar c\left(\frac{dF}{dr} - \frac{k}{r}F\right) = -(E - V - mc^2)G, \tag{6.18a}$$
$$\hbar c\left(\frac{dG}{dr} + \frac{k}{r}G\right) = (E - V + mc^2)F. \tag{6.18b}$$

Solving the Radial Equations

To solve for the radial functions we introduce the short-hand notation [9, p.172]

$$\begin{aligned}\alpha_1 = (mc^2 + E)/\hbar c, \quad \alpha_2 = (mc^2 - E)/\hbar c, \quad \sigma = \sqrt{\alpha_1\alpha_2},\\ \rho = r\sigma, \quad \gamma = (Ze^2/4\pi\hbar c) = Z\alpha,\end{aligned} \tag{6.19}$$

and substitute into (6.18) to acquire

$$\left(\frac{d}{d\rho} + \frac{k}{\rho}\right)F - \left(\frac{\alpha_2}{\sigma} - \frac{\gamma}{\rho}\right)G = 0, \tag{6.20a}$$
$$\left(\frac{d}{d\rho} - \frac{k}{\rho}\right)G - \left(\frac{\alpha_1}{\sigma} + \frac{\gamma}{\rho}\right)F = 0. \tag{6.20b}$$

This couple of equations can be solved by applying the differential operator $d/d\rho$ in one of the equations and then substitute the result into the other [9, p.173]. It can then be found that for large ρ's

$$\frac{d^2F}{d\rho^2} = F, \qquad \frac{d^2G}{d\rho^2} = G \tag{6.21}$$

In other words the solutions will behave as exponential functions $e^{\pm\rho}$, with $e^{-\rho}$ being the normalisable solution. This motivates the power series

$$F = e^{-\rho} f(\rho) \equiv e^{-\rho}\rho^s \sum_{m=0} a_m \rho^m, \qquad G = e^{-\rho} g(\rho) \equiv e^{-\rho}\rho^s \sum_{m=0} b_m \rho^m, \tag{6.22}$$

which when substituted into (6.20) turn the equations into

$$f' - f + \frac{kf}{\rho} - \left(\frac{\alpha_2}{\sigma} - \frac{\gamma}{\rho}\right) g = 0, \tag{6.23a}$$

$$g' - g - \frac{kg}{\rho} - \left(\frac{\alpha_1}{\sigma} + \frac{\gamma}{\rho}\right) f = 0. \tag{6.23b}$$

Further conclusions about the power series can be drawn by looking at the coefficients in front of $\rho^{s+\nu-1}$ after the series have been inserted into (6.23). The coefficients are

$$(s+\nu+k)b_\nu - b_{\nu-1} + \gamma a_\nu - \frac{\alpha_2}{\sigma} a_{\nu-1} = 0, \tag{6.24a}$$

$$(s+\nu-k)a_\nu - a_{\nu-1} - \gamma b_\nu - \frac{\alpha_1}{\sigma} b_{\nu-1} = 0. \tag{6.24b}$$

which gives a recursive relationship between them. Setting $\nu = 0$ we get

$$\begin{aligned} (s+k)b_0 + \gamma a_0 &= 0, \\ (s-k)a_0 - \gamma b_0 &= 0, \end{aligned} \tag{6.25}$$

which has non-zero solutions if the determinant is zero. Which is to say that

$$s = \pm\sqrt{k^2 - \gamma^2}. \tag{6.26}$$

To determine which sign is most suitable we take into consideration that the functions need to be normalisable. This can only be achieved if we take the positive root of (6.26).

Using the recursive relationship between the coefficients it can be shown that the series converge to $e^{2\rho}$ [8, p.486]. Because of this the series need to be finite i.e. there must exist an n' such that

$$a_{n'+1} = b_{n'+1} = 0, \qquad a_{n'} \neq 0, b_{n'} \neq 0. \tag{6.27}$$

Setting $\nu = n' + 1$ in (6.24) with the termination assumption in (6.27) gives the ratio

$$\frac{a_{n'}}{b_{n'}} = -\sqrt{\frac{\alpha_2}{\alpha_1}}. \tag{6.28}$$

If we multiply (6.24a) by σ and (6.24b) by α_2, take the difference between the two and set $\nu = n'$ we can obtain an equation involving just $a_{n'}$ and $b_{n'}$. This gives

$$2\sigma(s+n') = \gamma(\alpha_1 - \alpha_2) = \gamma E \tag{6.29}$$

and substituting back everything we defined in (6.19) we get

$$\sqrt{(mc^2)^2 - E^2}(s+n') = \gamma E \tag{6.30}$$

and finally

$$E = \frac{mc^2}{\sqrt{1+\frac{\gamma^2}{(s+n')^2}}} = \frac{mc^2}{\sqrt{1+\frac{Z^2\alpha^2}{(n'+\sqrt{(j+\frac{1}{2})^2 - Z^2\alpha^2})^2}}}. \tag{6.31}$$

This is the energy levels for an electron in an hydrogen atom solved exactly, no simplifying assumptions were made along the way. In this form it might be hard to draw any conclusions, but it is easily checked that there are no variables in (6.31) that we aren't familiar with. We can make a definition for the principal quantum number,

$$n \equiv n' + |k| = n' + j + \frac{1}{2}, \tag{6.32}$$

and then write

$$E = mc^2\left[1 + \left(\frac{Z\alpha}{n - (j+\frac{1}{2}) + \sqrt{(j+\frac{1}{2})^2 - (Z\alpha)^2}}\right)^2\right]^{-1/2} \tag{6.33}$$

which makes it more clear that n corresponds to the n from non-relativistic quantum mechanics. Eq. (6.33) is the exact fine-structure formula for hydrogen and if expanded in terms of α to order α^4 we obtain the energy levels with fine structure corrections as obtained from perturbation theory,

$$E = mc^2 - \frac{E_1}{n^2}\left[1 + \frac{\alpha^2}{n^2}\left(\frac{n}{j} + \frac{1}{2} - \frac{3}{4}\right)\right], \tag{6.34}$$

where E_1 is the ground state energy of hydrogen as calculated with Schrödinger's equation [6, p.275-276]. There is also the term corresponding to the rest energy, which is usually not considered in the non-relativistic case. Except for this rest energy term, this is of course the sum of the energy levels and correction terms obtained for the fine structure [6, p.270-274]

$$E = E_n + E_{fs}, \tag{6.35a}$$

$$E_{fs} = \frac{(E_n)^2}{2mc^2}\left(3 - \frac{4n}{j+1/2}\right) = E_r + E_{so}, \tag{6.35b}$$

$$E_{so} = \frac{(E_n)^2}{mc^2}\left[\frac{n[j(j+1) - l(l+1) - 3/4]}{l(l+1/2)(l+1)}\right], \tag{6.35c}$$

$$E_r = -\frac{(E_n)^2}{2mc^2}\left[\frac{4n}{l+1/2} - 3\right], \tag{6.35d}$$

where E_n is the energy levels $-E_1/n^2$. E_{so} is the correction from spin-orbit coupling and E_r is the relativistic correction.[6] To obtain the corresponding correction operators to the Hamiltonian we will have to take a different approach, this is done in the next section.

[6] One could also introduce the energy correction from the Darwin term to deal with the difficulties arising with $l = 0$ for E_r and E_{so}. The sum of the two correction is however correct for any l so this will be omitted [6, p.274].

7 Foldy-Wouthuysen Transformation

The solutions to the Dirac spinor ψ we have obtained in the earlier sections are built up of components that correspond to positive energy solutions and to negative energy solutions. To get a meaningful classical limit of the Dirac equation we could wish for it to converge to the Pauli theory of spin-$\frac{1}{2}$ particles, where the states are described by two-component wave functions.[7] One way of doing this is making use of the fact that in the non-relativistic limit two out of the four components of each spinor becomes negligible. If the small component equations are solved and substituted into the equations for the large components one approximately gets the Pauli spin equations [5, p.29]. However, according to Foldy and Wouthuysen, with this method there are problems. All four spinor components still need to be used when calculating expectation values to the order $\frac{v^2}{c^2} \approx \alpha^2$ and there are difficulties concerning the operators used in the two theories. In Dirac Pauli representation the velocity operator, $\dot{\boldsymbol{x}} = c\boldsymbol{\alpha}$, has eigenvalues $\pm c$ (since the components of $\boldsymbol{\alpha}$ has eigenvalues ± 1).[8] First of all this does not agree with the physical world. Secondly, it does not agree with the velocity operator of Pauli theory where the operator is $\boldsymbol{p}/mc^2$ and has eigenvalues ranging from $-\infty$ to ∞ [3, p.3]. In the Dirac theory the components of velocity do not even commute, while as in Pauli theory they do.

Foldy and Wouthuysen showed in their famous paper [5] that for a particle behaving according to the Dirac equation there exist another representation in which the positive and negative energy states are described by separate two-component wave functions, true for both relativistic and non-relativistic energies. For the free electron this representation can be obtained exactly and for the electromagnetic potential it can be done in a series to any order. Foldy and Wouthuysen also showed that in this representation there exists a position operator whose time derivative can be interpreted as a velocity operator which in the non-relativistic limit agrees with the one found in the Pauli representation [5, p.30]. Further, in representation that we have studied, the Dirac-Pauli representation, we've seen that $\boldsymbol{L}$ and $\boldsymbol{S}$ are not constants of motion. This required us to work with the total angular momentum operator $\boldsymbol{J}$. Foldy and Wouthuysen found that in their representation both the orbital angular momentum and the spin are constants of motion [3, p.3].

[7] This was one of the problems with the Dirac Equation according to Foldy and Wouthuysen in their paper *On the Dirac Theory of Spin 1/2 Particles and Its Non-Relativistic Limit* [5].

[8] A short derivation of this operator can be found in Appendix B.

Free particle

Adapting to natural units as in the original paper we set $\hbar = c = 1$. The transformation is presented following [5, p.30-31]. The canonical (unitary) transformation that achieves the required decoupling is

$$\psi' = e^{iS}\psi, \tag{7.1}$$

where S is not necessarily time-independent. The transformation of the Hamiltonian can be derived from the Dirac equation $E\psi = H\psi$. Transforming the left-hand side we get

$$E\psi = i\partial_t\psi = i\partial_t(e^{-iS}\psi') = i(\partial_t e^{-iS})\psi' + ie^{-iS}(\partial_t\psi'). \tag{7.2}$$

Transforming the right-hand side we get

$$H\psi = He^{-iS}\psi' \tag{7.3}$$

Setting (7.2) equal to (7.3), multiplying on the left with e^{iS} and rearranging we find that

$$i\partial_t\psi' = e^{iS}\big(He^{-iS} - (\partial_t e^{-iS})\big)\psi' \tag{7.4}$$

The transformed Hamiltonian can then be written as

$$H' = e^{iS}(H - i\partial_t)e^{-iS}. \tag{7.5}$$

Foldy and Wouthuysen proceed to define

$$S = -(i/2)\frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\tan^{-1}(p/m). \tag{7.6}$$

Now the choice of S may seem mysterious, especially with the trigonometric function. The explanation lies in the fact that the goal is to construct S in such a way that H' contains no operators that mix the large and small components of the Dirac spinor. The matrix operators that do this will be referred to as odd operators. For example the α and γ matrices (except γ^0) cause this mixing and are hence odd. Even matrices on the other hand do not cause this difficulty, like the $\beta = \gamma^0$ and $\boldsymbol{\Sigma}$ matrices [9, p.181]. The S defined in (7.6) does the job, it removes all the the odd operators of the Hamiltonian.

A slightly more explanatory approach is taken by Drell & Bjorken in *Relativistic Quantum Mechanics* [1, p.45-48]. They define

$$S \equiv -i\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}\theta \quad \rightarrow \quad e^{iS} = e^{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}\theta}. \tag{7.7}$$

If we expand the exponential function in terms of θ we obtain

$$e^{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}\theta} = 1 + (\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})\theta + (\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})^2\frac{\theta^2}{2} + (\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})^3\frac{\theta^3}{6} + ..., \tag{7.8}$$

which can be simplified using the relations

$$(\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 = \alpha^i\alpha^j p^i p^j = \boldsymbol{p}^2, \tag{7.9a}$$

$$(\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 = -\beta^2(\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 = -\boldsymbol{p}^2. \tag{7.9b}$$

We then get that

$$e^{iS} = 1 + (\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})\theta - \boldsymbol{p}^2\frac{\theta^2}{2} - \boldsymbol{p}^2(\beta\boldsymbol{\alpha}\cdot\boldsymbol{p})\frac{\theta^3}{6} + \boldsymbol{p}^4\frac{\theta^4}{24} + \dots \tag{7.10}$$

If we compare the result of (7.10) with the series of the trigonometric functions it is straightforward to verify that we have

$$e^{iS} = \cos(p\theta) + \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta) \tag{7.11}$$

Similarily we have for the inverse

$$e^{-iS} = \cos(p\theta) - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta). \tag{7.12}$$

Substituting these results into the Hamiltonian transformation (7.5) we obtain

$$H' = \left(\cos(p\theta) + \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta)\right)H\left(\cos(p\theta) - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta)\right),$$

since S is time-independent [9, p.182]. The free particle hamiltonian, $H = \boldsymbol{\alpha}\cdot\boldsymbol{p} + m\beta$, anticommutes with the factor in front of the $\sin(p\theta)$ term (see eq. (3.16)) so that

$$\begin{aligned} H' = \left(\cos(p\theta) + \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta)\right)H\left(\cos(p\theta) - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(p\theta)\right) \\ = H\left(e^{-2\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}\theta}\right) \\ = \boldsymbol{\alpha}\cdot\boldsymbol{p}\left(\cos(2p\theta) - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(2p\theta)\right) + \beta m\left(\cos(2p\theta) - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(2p\theta)\right) \end{aligned}$$

Performing the matrix multiplication and once again keeping in mind the commutation relations

$$\begin{aligned} H' &= \boldsymbol{\alpha}\cdot\boldsymbol{p}\cos(2p\theta) + \beta p\sin(2p\theta) + \beta m\cos(2p\theta) - m\frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}\sin(2p\theta) \\ &= \boldsymbol{\alpha}\cdot\boldsymbol{p}(\cos(2p\theta) - \frac{m}{p}\sin(2p\theta)) + \beta m(\cos(2p\theta) + \frac{\boldsymbol{p}}{m}\sin(2p\theta)). \end{aligned}$$

What is left is choosing θ in such a way that the odd operators of the expression vanishes. This motivates $\tan(2p\theta) = p/m$, or equivalently $\theta = (1/2p)\tan^{-1}(p/m)$ which turns S into the same one that Foldy and Wouthuysen defined in their paper (eq. 7.6). Substituting this we obtain the new Hamiltonian [1, p.48]

$$H' = \beta(m^2 + \boldsymbol{p}^2)^{1/2}. \tag{7.16}$$

This Hamiltonian, unlike the one from the Dirac representation is in agreement with the one from classical physics [3, p.3]. Notice that from the definition of β, eq. (3.15), we get two uncoupled equations. Two with a positive sign in front and two with a negative sign.

Electromagnetic Field

In the case of an electromagnetic field the transformation is more complicated. The Hamiltonian is then given by

$$H = \boldsymbol{\alpha} \cdot (\boldsymbol{p} - e\boldsymbol{A}) + \beta m + e\Phi \tag{7.17}$$

where $\boldsymbol{A}$ is the vector potential and Ψ is the scalar potential. To solve this we write the Hamiltonian in an simplified manner as

$$H = \beta m + \mathcal{E} + \mathcal{O} \tag{7.18}$$

where $\mathcal{O} = \boldsymbol{\alpha} \cdot (\boldsymbol{p} - e\boldsymbol{A})$ is the even term and $\mathcal{E} = e\Phi$ is the odd term [9, p.183-184]. The terms carry the properties

$$\beta\mathcal{E} = \mathcal{E}\beta, \tag{7.19a}$$
$$\beta\mathcal{O} = -\mathcal{O}\beta. \tag{7.19b}$$

Following the steps of Bjorken and Drell [1, p.48-49], we expand the transformation equation for the Hamiltonian with the help of the Baker-Hausdorff formula.[9] The following expression is obtained to n'th order

$$\begin{aligned} H' = e^{iS}(H - i\partial_t)e^{-iS} = e^{iS}He^{-iS} - e^{iS}\dot{S}e^{-iS} = \\ H + i[S,H] + \frac{i^2}{2!}[S,[S,H]] + \frac{i^3}{6}[S,[S,[S,H]]]] + ... + \frac{i^n}{n!}[S,[S,...[S,H]...]] + \\ ... - \dot{S} - \frac{i}{2}[S,\dot{S}] - \frac{i^3}{6}[S,[S,[S,\dot{S}]]]] - ... - \frac{i^n}{n!}[S,[S,...[S,\dot{S}]...]] - ... \end{aligned}$$

To get a better feeling for this expression we can see that for small p/m, i.e. in the non-relativitic limit, we have from eq. (7.6) that

$$iS \approx (1/2)\frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{p}(p/m) = \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{2m} \tag{7.21}$$

which is the same to say as $iS \propto \frac{1}{m}$ and thus we interpret the long expression in (7.20) as a series in $\frac{1}{m}$ [9, p.184]. Following the approximation of (7.21), where S contains the odd term of the free particle Hamiltonian, we make an ansatz that

$$S = -i\beta\frac{\boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A})}{2m} = -i\beta\mathcal{O}/2m, \tag{7.22}$$

which also makes S an odd term. In the non-relativistic limit we terminate the series so that it takes the form

$$\begin{aligned} H' = H + i[S,H] + \frac{i^2}{2!}[S,[S,H]] + \frac{i^3}{6}[S,[S,[S,H]]]] + \\ \frac{i^4}{24}[S,[S,[S,[S,H]]]] - \dot{S} - \frac{i}{2}[S,\dot{S}] - \frac{i^3}{6}[S,[S,[S,\dot{S}]]]] \end{aligned}$$

[9] $e^A B e^{-A} = B + [A,B] + \frac{1}{2}[A,[A,B]] + ... + \frac{1}{n!}[A,[A,...[A,B]...]] + ...$ [9, p.184].

All we need to get our result are odd operators of order lower than $(\frac{1}{m})^2$ and even operators of order lower than $(\frac{1}{m})^3$, and hence all the other terms can be neglected [9, p.184]. It should be mentioned that the product of two odd and two even operators gives an odd operator, while mixing even and odd gives an odd operator [5, p.35]. In other words, if there are operators $\frac{1}{m^4}\mathcal{O}^2$, even, and $\frac{1}{m^3}\mathcal{E}\mathcal{O}$, odd, they will be neglected. It may also be good to remind ourselves that that β is an even operator. What is left is now to evaluate the terms of (7.23). From [1, p.50] we obtain

$$i[S,H] = -\mathcal{O} + \frac{\beta}{2m}[\mathcal{O},\mathcal{E}] + \frac{1}{m}\beta\mathcal{O}^2, \tag{7.24a}$$

$$\frac{i^2}{2}[S,[S,H]] = -\frac{\beta\mathcal{O}^2}{2m} - \frac{1}{8m^2}[\mathcal{O},[\mathcal{O},\mathcal{E}]] - \frac{1}{2m^2}\mathcal{O}^3, \tag{7.24b}$$

$$\frac{i^3}{3!}[S,[S,[S,H]]] = \frac{\mathcal{O}^3}{6m^2} - \frac{1}{6m^3}\beta\mathcal{O}^4, \tag{7.24c}$$

$$\frac{i^4}{4!}[S,[S,[S,[S,H]]]] = \beta\frac{\beta\mathcal{O}^4}{24m^3} \tag{7.24d}$$

which also can be calculated using the properties of (7.19), neglecting operators with a factor $\frac{1}{m}$ larger than sufficient. From these equations, though they look messy, two important conclusions can be drawn. First, since we have $H' = H + i[S,H] + ...$ where $H = \beta m + \mathcal{E} + \mathcal{O}$ the transformation has removed the initial odd operator we had from the Hamiltonian. We have however, as the equations above show, introduced several new odd operators. Even though it might not be completely obvious, there are in fact no odd operators of order $(\frac{1}{m})^0$ left [4, p.35]. The final terms are [1, p.50]

$$\dot{S} = -i\frac{i\beta\dot{\mathcal{O}}}{2m}, \tag{7.25a}$$

$$-\frac{i}{2}[S,\dot{S}] = -\frac{i}{8m^2}[\mathcal{O},\dot{\mathcal{O}}]. \tag{7.25b}$$

The crucial part is that we've removed the odd operator of order $(\frac{1}{m})^0$ and now have odd operators of at most order $\frac{1}{m}$. We still want to remove these and the way to do this is to apply the same transformation in a recursive manner with a new S, until they agree with the criteria under which we neglect odd operators. To do this we first gather the odd and even operators and define $\mathcal{E}'$ and $\mathcal{O}'$ [1, p.50]

$$\mathcal{E}' = \mathcal{E} + \beta\left(\frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) - \frac{1}{8m^2}[\mathcal{O},[\mathcal{O},\mathcal{E}]] - \frac{1}{8m^2}[\mathcal{O},\dot{\mathcal{O}}], \tag{7.26a}$$

$$\mathcal{O}' = \frac{\beta}{2m}[\mathcal{O},\mathcal{E}] - \frac{\mathcal{O}^3}{3m^2} + \frac{i\beta\dot{\mathcal{O}}}{2m}. \tag{7.26b}$$

which then leads us to define the new Hamiltonian

$$H' \equiv \beta m + \mathcal{E}' + \mathcal{O}'. \tag{7.27}$$

In analogy with how we defined S to remove the operators of order $(\frac{1}{m})^0$ we will define

$$S' = \frac{i\beta}{2m}\mathcal{O}', \tag{7.28}$$

which now is of order $(\frac{1}{m})^2$. We go through the same procedure again, neglecting terms not fulfilling the requirements. With this iteration we obtain [9, p.185]

$$H'' = \beta m + \mathcal{E}' + \frac{\beta}{2m}[\mathcal{O}', \mathcal{E}'] + \frac{i\beta\dot{\mathcal{O}}'}{2m} \tag{7.29}$$

as our transformed Hamiltonian and this time the odd operators are of at most order $\frac{1}{m^2}$. Performing the transformation one more time with $S'' = i\beta\mathcal{O}''/2m$ we obtain

$$H''' = \beta m + \mathcal{E}' \tag{7.30}$$

since now the odd operator $\mathcal{O}''$ is of order $(\frac{1}{m})^3$ and can be neglected [9, p.186]. What is left is to evaluate the even operators that remain

$$\beta m + \mathcal{E}' = \beta m + \mathcal{E} + \beta\left(\frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) - \frac{1}{8m^2}\left[\mathcal{O}, [\mathcal{O}, \mathcal{E}] + i\dot{\mathcal{O}}\right], \tag{7.31}$$

which of course requires combining the terms in correct ways. Omitting the calculations, see for example [9, p.186], we obtain

$$\begin{aligned} H''' = \beta\left(m + \frac{(\boldsymbol{p} - e\boldsymbol{A})^2}{2m} - \frac{1}{8m^3}[(\boldsymbol{p} - e\boldsymbol{A})^2 - e\boldsymbol{\Sigma}\cdot\boldsymbol{B}]^2\right) + e\Phi \\ - \frac{e}{2m}\beta\boldsymbol{\Sigma}\cdot\boldsymbol{B} - \frac{ie}{8m^2}\boldsymbol{\Sigma}\cdot\text{curl }\boldsymbol{E} \\ - \frac{e}{4m^2}\boldsymbol{\Sigma}\cdot\boldsymbol{E}\times(\boldsymbol{p} - e\boldsymbol{A}) - \frac{e}{8m^2}\text{div }\boldsymbol{E}. \end{aligned} \tag{7.32}$$

This Hamiltonian completely decouples the upper two components from the lower two components of the bispinor since there are no odd operators. Notice for example that because of the β (diagonal matrix with first two diagonal components $(+1)$ and the two second (-1)) the m term (mc^2 in SI-units) will be negative for the first two components of the bispinor and negative for the second two components. This supports the interpretation that the first components of the bispinor are positive energy solutions and the second two negative energy solutions. Based on [9, p.187] we let the Hamiltonian operate on a spinor $\psi = \begin{bmatrix}\psi_1 \\ \psi_2\end{bmatrix}$ and get for the component ψ_1

$$\begin{aligned} i\partial_t\psi_1 = \Bigg[m + e\Phi + \frac{1}{2m}(\boldsymbol{p} - e\boldsymbol{A})^2 - \frac{e}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} - \frac{\boldsymbol{p}^4}{8m^3} \\ - \frac{e}{4m^2}\boldsymbol{\Sigma}\cdot\boldsymbol{E}\times(\boldsymbol{p} - e\boldsymbol{A}) - \frac{e}{8m^2}\text{div }\boldsymbol{E}\Bigg]\psi_1. \end{aligned} \tag{7.33}$$

For the hydrogen atom we have a spherical constant potential $\Phi(r)$, which leads to $\boldsymbol{E} = -\nabla\Phi(r) = -\frac{1}{r}\frac{\partial\Phi}{\partial r}\boldsymbol{r}$ and $\boldsymbol{A} = 0$. This allows for the simplification [9, p.187]

$$\boldsymbol{\Sigma}\cdot\boldsymbol{E}\times\boldsymbol{p} = -\frac{1}{r}\frac{\partial\Phi}{\partial r}\boldsymbol{\Sigma}\cdot\boldsymbol{r}\times\boldsymbol{p} = -\frac{1}{r}\frac{\partial\Phi}{\partial r}\boldsymbol{\Sigma}\cdot\boldsymbol{L}. \tag{7.34}$$

Substituting (7.34) into (7.33) we can identify the following operators.

$$\begin{aligned}
H_0 &= m && \text{(rest energy)}\\
H_1 &= e\Phi(r) && \text{(potential energy)}\\
H_3 &= \frac{1}{2m}(\boldsymbol{p} - e\boldsymbol{A})^2 && \text{(kinetic energy)}\\
H_\mu &= \frac{e}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} && \text{(coupling of magnetic moment)}\\
H_r &= -\frac{(\boldsymbol{p}^2)^2}{8m^3} && \text{(relativistic correction)}\\
H_{so} &= \frac{e}{4m^2}\frac{1}{r}\frac{\partial\Phi}{\partial r}\boldsymbol{\Sigma}\cdot\boldsymbol{L} && \text{(spin-orbit correction)}\\
H_d &= -\frac{e}{8m^2}\nabla^2\Phi(r) && \text{(Darwin term)}
\end{aligned} \tag{7.35}$$

The fourth term, the coupling of the magnetic moment, usually written as $\boldsymbol{\mu}\cdot\boldsymbol{B}$ where $\boldsymbol{\mu} \equiv -g\frac{e}{2m}\frac{1}{2}\boldsymbol{\sigma}$ [6, p.271-272], implies that the Landé g-factor for the electron is $g = 2$. This agrees with experiments and is a further indication that the Dirac equation is a powerful tool.

8 Conclusions

We have, by applying two different methods, been able to obtain some of the most interesting corrections to the Schrödinger theory. The Taylor expansion of the exact energy expression gave us the energy correction terms and the transformation due to Foldy and Wouthuysen gave us the corresponding correction terms to the Hamiltonian. There are more corrections that can be found and are needed to explain phenomena such as hyperfine splitting. This would however require that the electron's interaction with the nucleus is taken into consideration, for example introducing the spin of the nucleus. There is also Lamb shift, which is not predicted by the Dirac equation. The energy correction due to the Lamb shift is slightly smaller, being of order α^5 which was also outside the scope of this text. The discovery of the Dirac equation is seen as one of the greatest achievements of modern physics. Dirac had his own idea of how to interpret the negative energy solutions. He presented his famous 'Dirac sea' where all the negative energy states were occupied by electrons, protected by Pauli's exclusion principle. In this model the vacuum around us is a sea of infinite negative energy and the positron can be explained as a hole moving through the sea. Today we are more comfortable with the idea of anti-matter, which quantum field theory can explain in a more rigorous way.

References

[1] James. D. Bjorken and Sidney D. Drell. *Relativistic Quantum Mechanics.* McGraw-Hill, Inc., 1964.

[2] Jim Branson. *Dirac Equation.* URL: `http://quantummechanics.ucsd.edu/ph130a/130_notes/node477.html`.

[3] John P. Costella and Bruce H. J. McKellar. "The Foldy-Wouthuysen transformation". In: *Am.J.Phys.* 63. 1119. (1995). URL: `arXiv:hep-ph/9503416`.

[4] P. A. M. Dirac. "On the Dirac Theory of Spin 1/2 Particles and Its Non-Relativistic Limit". In: *Proc. R. Soc.* 117.778 (1928), pp. 610–624. DOI: `10.1098/rspa.1928.0023`.

[5] Leslie L. Foldy and Siegfried A. Wouthuysen. "On the Dirac Theory of Spin 1/2 Particles and Its Non-Relativistic Limit". In: *Physical Review* 78.1 (1950), pp. 29–36. URL: `http://einstein.drexel.edu/~bob/Quantum_Papers/Foldy-Wouthuysen.pdf`.

[6] David J. Griffiths. *Introduction to Quantum Mechanics.* Pearson Prentice Hall. 2nd Edition., 2005. ISBN: 0-13-191175-9.

[7] J.J. Sakurai. *Advanced Quantum Mechanics.* Addison-Wesley Publishing Company, Inc. 11th Printing., 1987. ISBN: 0-201-06710-2.

[8] Leonard I. Schiff. *Quantum Mechanics.* McGraw-Hill, Inc. 3rd Edition., 1968. ISBN: 0-201-06710-2.

[9] Franz Schwabl. *Quantenmechanik für Fortgeschrittene (QM II). (German) [Advanced Quantum Mechanics].* Springer-Verlag Berlin Heidelberg. 3rd Edition., 2000. ISBN: 3-540-67730-5.

Appendix A

Relativistic Notation

The cornerstone on which special relativity stands on is the concept of space-time. In space-time we specify an event by 4 coordinates,

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \equiv (ct, \boldsymbol{x}). \tag{8.1}$$

Observers moving relative to each other will not be able to agree on time and distance between two events but besides the speed of light there is one quantity they will agree on. Given that one observer measures the difference in space-time coordinates between two events to be (ct, x, y, z) and another observer that measures the same distance to be (ct', x', y', z') they will both agree on the quantity

$$s^2 \equiv c^2t^2 - x^2 - y^2 - z^2 = c^2(t'^2) - (x')^2 - (y')^2 - (z')^2. \tag{8.2}$$

If we are interested in infinitesimal distances in space-time we might as well write the same expression as

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{8.3}$$

Introducing Einstein's sum convention, where summation is performed over indices appearing both as super-scripts and sub-scripts, and the symmetric metric of special relativity

$$g_{\mu\nu} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{8.4}$$

we can write

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{8.5}$$

The metric $g_{\mu\nu}$ is its own inverse, so that multiplying it with a vector twice is equal to just multiplying with the identity matrix. For convenience we could denote the inverse as $g^{\mu\nu}$ to go along with Einstein's sum convention. This will also be more clear if we define contravariant and covariant vectors. Given a position 4-vector x^μ we have

$$g_{\mu\nu} x^\mu = g_{\mu\nu}(x^0, \boldsymbol{x}) = (x^0, -\boldsymbol{x}) \equiv x_\mu \tag{8.6}$$

A vector denoted by a subscript will be called a covariant vector while as a contravariant vector will be denoted by a superscript. Similarily we get

$$g^{\mu\nu} x_\mu = g^{\mu\nu}(x^0, -\boldsymbol{x}) = (x^0, \boldsymbol{x}) \equiv x^\mu. \tag{8.7}$$

To get expressions for the energy and momentum of special relativity we will imagine a situation where a particle is moving relative to a frame of reference while it is at rest according to another frame of reference. The space-time difference between observing the particle twice will in the first frame of reference lead to

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = cdt^2\left(1 - \frac{v^2}{c^2}\right). \tag{8.8}$$

In the second reference frame, where the particle is at rest, we obtain

$$ds^2 = c^2(dt')^2. \tag{8.9}$$

We call the time $\tau \equiv t'$ the proper time, it is the time that the moving particle 'experiences'. The relationship between the different t's is thus

$$d\tau = dt\sqrt{1 - \frac{v^2}{c^2}}. \tag{8.10}$$

Using the proper time we define the 4-velocity to be

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\left(c, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dx}{dt}\right) \equiv \gamma(c, \boldsymbol{v}). \tag{8.11}$$

The 4-momentum is defined in the straightforward way

$$p^\mu \equiv mu^\mu = \gamma(mc, m\boldsymbol{v}) \equiv (\gamma mc, \boldsymbol{p}). \tag{8.12}$$

Kinetic energy in special relativity is given by

$$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{8.13}$$

and hence 4-momentum can also be written as

$$p^\mu = (E/c, \boldsymbol{p}). \tag{8.14}$$

Appendix B

Heisenberg Picture

An alternative representation of Quantum Mechanics is the so called Heisenberg picture. In Schrödinger's formulation the state vectors carry with them a time dependence. In the Heisenberg picture the state vectors themselves do not depend on time while the operators do. This gives an alternate representation that can be very useful, especially when investigating the time dependencies of expectation values.

To start with we will look at the Schrödinger picture where the Schrödinger equation can be written

$$i\hbar \frac{d}{dt}|\psi, t\rangle = \hat{H}|\psi, t\rangle. \tag{8.15}$$

If there is no time-dependence in the Hamiltonian the time-dependent state vector $|\psi, t\rangle$ is correlated to an initial state by a time dependent factor [8, p.169],

$$|\psi, t\rangle = e^{-i\hat{H}(t-t_0)/\hbar}|\psi, t_0\rangle \equiv \hat{U}|\psi, t_0\rangle. \tag{8.16}$$

The expectation value of some operator $\hat{A}$ can then be written

$$\begin{aligned}\langle A\rangle = \langle\psi, t|A|\psi, t\rangle = \langle\psi, t_0|\hat{U}^\dagger A\hat{U}|\psi, t_0\rangle \\ = \langle\psi, t_0|e^{i\hat{H}(t-t_0)/\hbar}Ae^{-i\hat{H}(t-t_0)/\hbar}|\psi, t_0\rangle,\end{aligned} \tag{8.17}$$

In the above expression we have used that [8, p.169]

$$\hat{U}^\dagger = \hat{U}^{-1} = e^{i\hat{H}(t-t_0)/\hbar}, \tag{8.18}$$

which further makes $\hat{U}$ a unitary operator ($\hat{U}^{-1}\hat{U} = \hat{U}^\dagger\hat{U} = 1$). In the Heisenberg picture the expectation values are calculated with the same expression (8.17) but instead state vectors and the operators are defined differently. To compare the two we will introduce a sub-indexes to keep track of the two pictures. The operator $\hat{A}$ in the two pictures is related by

$$\hat{A}_H = \hat{U}^\dagger\hat{A}_S\hat{U} \tag{8.19}$$

and the state vectors [7, p.113]

$$|\psi, t\rangle_H = \hat{U}^{-1}|\psi, t\rangle_S = |\psi, t_0\rangle_S. \tag{8.20a}$$

We can investigate the time dependence of the operator by differentiating. We obtain

$$\frac{d\hat{A}_H}{dt} = \frac{\partial\hat{U}^\dagger}{\partial t}\hat{A}_S\hat{U} + \hat{U}^\dagger\frac{\partial\hat{A}_S}{\partial t}\hat{U} + \hat{U}^\dagger\hat{A}_S\frac{\partial\hat{U}}{\partial t} \tag{8.21a}$$

$$= \big(\frac{i}{\hbar}\hat{H}\big)\hat{U}^\dagger\hat{A}_S\hat{U} + \hat{U}^\dagger\hat{A}_S\big(\frac{-i}{\hbar}\hat{H}\big)\hat{U} + \hat{U}^\dagger\frac{\partial\hat{A}_S}{\partial t}\hat{U}, \tag{8.21b}$$

which can be simplified into

$$\frac{d\hat{A}_H}{dt} = \frac{i}{\hbar}\hat{H}\hat{A}_H - \frac{i}{\hbar}\hat{A}_H\hat{H}, \tag{8.22a}$$

since $(\partial\hat{A}_S/\partial t) = 0$ and $[\hat{H},\hat{U}] = 0$. The commutator can be motivated by the fact that $\hat{U}$ can be written in a series of $\hat{H}$ [8, p.169]. What is left is just the commutator of $\hat{A}_H$ and $\hat{H}$ with an extra factor,

$$\frac{d\hat{A}_H}{dt} = \frac{1}{i\hbar}[\hat{A}_H,\hat{H}]. \tag{8.23}$$

With eq. (8.23) we can then investigate whether observables are conserved, so called *constants of motion*. Based on the derivations made by J.J. Sakurai [7, p.113-114] we study the operators $\boldsymbol{L}$ and $\boldsymbol{S}$. For a free particle we have the Hamiltonian, simplified using Einsteins sum convention, $H = c\alpha_j p_j + \beta mc^2$. The orbital angular momentum operator $\boldsymbol{L} \equiv \boldsymbol{x} \times \boldsymbol{p}$ is not a constant of motion since

$$\begin{aligned}\frac{d\hat{L}_1}{dt} = \frac{1}{i\hbar}[L_1,\hat{H}] = \frac{1}{i\hbar}[x_2p_3 - x_3p_2, c\alpha_j p_j + \beta mc^2]\\ = \frac{1}{i\hbar}[x_2p_3 - x_3p_2, c\alpha_j p_j]\\ = \frac{1}{i\hbar}\big([x_2p_3, c\alpha_j p_j] - [x_3p_2, c\alpha_j p_j]\big)\\ = c(\alpha_3 p_2 - \alpha_2 p_3)\end{aligned} \tag{8.24}$$

and the same applies to the other components. In the above calculation we have used that $p_j \equiv -i\hbar(\partial/\partial x_j)$ and $p_j p_k$ acting on a state vector is 0 for all spacial combinations j,k. Hence we have

$$\frac{d\boldsymbol{L}}{dt} = c(\boldsymbol{\alpha} \times \boldsymbol{p}) \neq 0. \tag{8.25}$$

We can show that the spin-operator $\boldsymbol{S} \equiv \frac{\hbar}{2}\boldsymbol{\Sigma}$ is not a constant of motion by using the identities of (5.11) and keeping in mind that swapping two gamma matrices of different index changes the sign of the product

$$\begin{aligned}\frac{\partial S_1}{\partial t} = \frac{1}{i\hbar}\frac{\hbar}{2}[\Sigma_1,\hat{H}] = \frac{1}{2i}[\Sigma_1, c\alpha_j p_j + \beta mc^2]\\ = \frac{1}{2i}[-i\gamma^2\gamma^3, c\gamma^0\gamma^j p_j]\\ = \frac{1}{2}cp_j\gamma^0(-\gamma^2\gamma^3\gamma^j + \gamma^j\gamma^2\gamma^3).\end{aligned} \tag{8.26}$$

If $j = 1$ the result is zero and non-zero otherwise. We are left with

$$\begin{aligned}\frac{\partial S_1}{\partial t} = c\gamma^0\frac{1}{2}(2p_2\gamma^3 - 2p_3\gamma^2) = c(p_2\gamma^0\gamma^3 - p_3\gamma^0\gamma^2)\\ = c(p_2\alpha_3 - p_3\alpha_2).\end{aligned} \tag{8.27}$$

Just like the components of the orbital angular momentum operator the spin operator's components are not constants of motions. The other components are calculated in the same way and leads to the conclusion that

$$\frac{d\boldsymbol{S}}{dt} = -c(\boldsymbol{\alpha} \times \boldsymbol{p}). \tag{8.28}$$

The Heisenberg picture may also be used to find operators. We could find the velocity operator in the Dirac theory by differentiating the components of the position operator $\boldsymbol{x}$. Doing this we get

$$\dot{x}_k = \frac{dx}{dt} = \frac{1}{i\hbar}[x_k, c\alpha_j p_j] = \frac{1}{i\hbar}c\alpha_j(x_k p_j - p_j x_k). \tag{8.29}$$

and from non-relativistic quantum mechanics we know that the last factor is just the canonical commutation relation which is equal to $i\hbar\delta_{kj}$ [6, p.133]. The result is simply

$$\dot{\boldsymbol{x}} = c\boldsymbol{\alpha}. \tag{8.30}$$