UNIVERSITY OF BERGEN

Department of Physics and Technology

MASTERS THESIS

Laser-atom Interactions in the Dirac Equation for Hydrogen

Author: Morten Tysse

Supervisor: Morten Førre

June 1, 2018

Acknowledgment

I would like to offer my sincerest gratitude to the atomic physics and optics group at the University of Bergen, for both social experiences and scientific insight. It has been an interesting couple of years, especially thanks to my supervisor Morten Førre, who has offered great support and extraordinary professional guidance. I would also like to thank Andreas Skeidsvoll, Ingunn Koren Rossland and Markus Lysne for successful cooperation and many fruitful discussions.

Abstract

The advancements of laser technology make it possible to produce high-energy lasers. It is of interest in many technologies, especially optics, to study how these high-energy lasers interact with atomic or molecular systems. The interaction needs to be modelled in a relativistic manner, due to the high-energy pulse. To get a basic understanding of the dynamics of the interaction, the field of atomic physics often look at the simplest case, hydrogenlike atoms.

This thesis models hydrogenlike atoms exposed to high-energy pulses in the Dirac equation and proposes numerical solutions. The high-energy pulse is modelled as a classical wave and is included in the Dirac equation by using the method of minimal coupling. The numerical solutions are within the dipole approximations, as the spatial dependence of the pulse is neglected and thus does not have a magnetic field. The dipole approximation reduces the validity of results with truly high-energy pulses, so it is suggested to go beyond the dipole approximation to work with such pulses.

The equation is first solved time-independently and then propagated in time by using the lowest order of the Magnus expansion and Krylov subspace methods. Differential probability distributions for the positive pseudo continuum and probabilities for photoionization in the Dirac equation is compared with its non-relativistic analogue, the Schrödinger equation.

Contents

Chapter 1

Introduction

1.1 Background

Due to advancements in laser technology, it is today possible to obtain pulses with high intensity. The University of Michigan has managed to make a laser, the HERCULES Petawatt Laser, that can produce intensities up to $2 \cdot 10^{22} \frac{W}{cm^2}$ [1]. The European X-ray Free Electron Laser in Hamburg has also shown great promise[2]. As intensity is energy transferred per second per area, high intensity implies high energy. These new high-energy lasers are of wide interest in many technologies, especially how they interact with atoms and molecules.[3]

Non-relativistic atoms and molecules are successfully described by quantum mechanics, but because the pulses are of high energy, the interactions are expected to be of relativistic character. In this regard, a theory that combines both quantum mechanics and special relativity is needed to model the interactions.

1.2 A Historical Peek Into Quantum Mechanics

Quantum mechanics is a theory that describes physical systems with few particles, like atoms and molecules. The theory can explain the structure and energies, among other physical properties, of these systems.

The theory of quantum mechanics was born throughout the 1800's with many different scientific discoveries. Already in 1872 it was suggested by Ludwig Boltzmann that the energy of a physical system, like atoms and molecules, only exist in discrete values; the energy was quan-

tized![4]. The photoelectric effect was discovered by Heinrich Hertz in 1887 (See figure (1.1)[5]. Light, or electromagnetic radiation in general, was previously believed to be continuous waves, but this was proven wrong by the photoelectric effect.

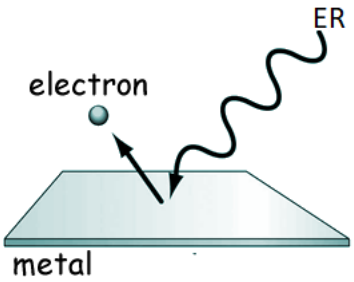


Figure 1.1: The photoelectric effect[6]. Electromagnetic radiation, later called photons, hits a metal object and emits an electron. It was shown by Heinrich Hertz that electrons are only emitted if the energy of the electromagnetic radiation, denoted ER, is above a certain threshold. This contradicted the assumption of electromagnetic radiation being continuous waves.

Albert Einstein explained the photoelectric effect in 1905 as he proposed that light is in fact made up of massless quantum particles, photons, and not continuous waves[7]. He had used Max Planck's idea from 1900, where it had been proposed that light could only be absorbed or emitted in discrete values called quanta[8].

This quantization of light emission and absorption increased the understanding of spectroscopy and spectral lines. The spectral lines of an atom or molecule must be consistent with the discrete energy states of the same atom or molecule. Niels Bohr managed to explain the spectral lines of hydrogen in 1913, as he proposed that the electron circuits the nucleus in discrete orbits and therefore only exists in discrete energy states[9]. See figure 1.2 .

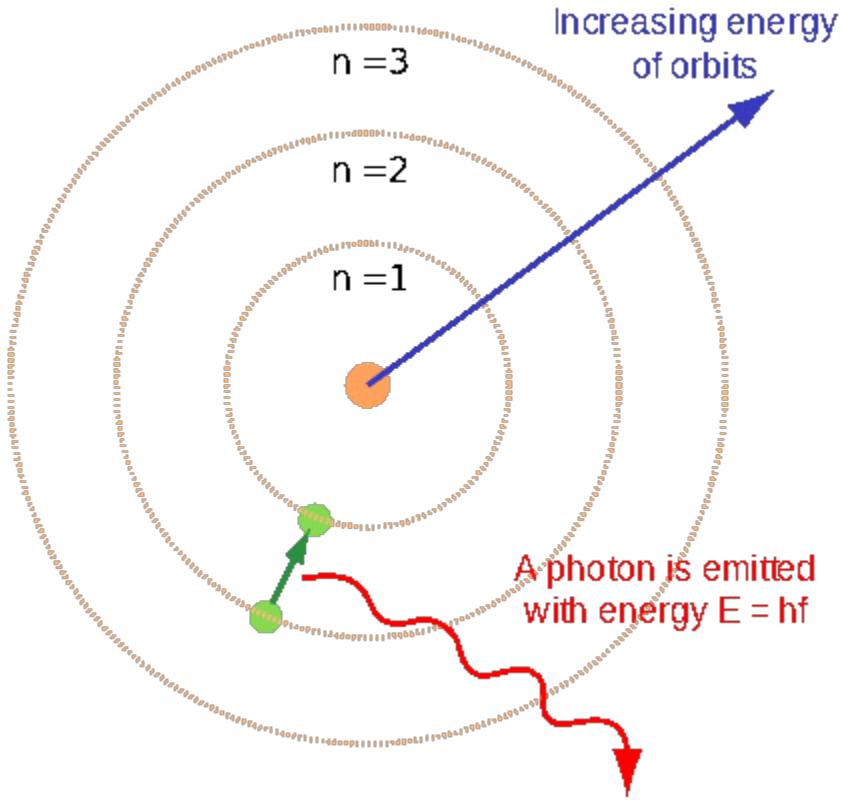


Figure 1.2: Illustration of Bohr's atom model[10]. The electron is only allowed to exist in certain orbits, numbered with main principal quantum numbers n. When the electron de-excites from a certain energy state to another, a photon is produced with energy corresponding to the energy difference of the two energy states.

It was now established that waves could behave as particles. Nine years later, in 1924, Louis de Broglie proposed that also particles could behave as waves[11]. This is called wave-particle duality. Wave-like properties are significant for atoms and molecules, but this is not the case for macroscopic objects due to their big mass. Nanomaterials have often small mass and therefore exhibit wave-like properties.

In 1925 Schrödinger proposed an equation, the Schrödinger equation, which shows how a wave develops in time. As particles act like waves according to the wave-particle duality, the Schrödinger equation was interpreted as particles evolving in time[12].

Max Born interpreted the Schrödinger equation as an equation that predicts the probability of a measurement to yield certain results. Quantum mechanics is therefore indeterministic. Another famous principle from quantum mechanics is the uncertainty principle, which was presented by Werner Heisenberg in 1927[13]. It states that certain physical properties of a particle or a physical system cannot be precisely defined at the same time. This applies to position and

momentum, among others.

The Schrödinger equation is the basis for modern quantum mechanics as it is very useful for modelling atomic systems and the study of quantum effects and energies. The Schrödinger equation does have its shortcomings though. It is only valid in non-relativistic cases, where the velocity or the energy of the system is relatively low. It needs to be merged with special relativity to be valid in cases with high-energy or velocity.

1.3 A Historical Introduction to Special Relativity

Humans experience the concept of time as something that develops at constant rates at all places. Time was therefore commonly accepted as universally absolute until the late 1800's. Galileo Galilei's principle of relativity, where relative velocities are additive, was also accepted as true.

Questions concerning the speed of light gave rise to the theory of special relativity. At the start of the 1800's, it was believed that light propagates as transverse waves through a mystical medium called ether. James Clerk Maxwell completed his theory on electromagnetism in 1873, where he predicted the speed of light in this ether[14]. The Michelson-Morley experiment rejected the existence of ether in 1887[15].

Einstein published a paper in 1905 where he made the concept of a medium like ether unnecessary[16]. He proposed two postulates of special relativity; the first being that the speed of light is constant in all reference frames, whereas the second states that the physical laws are equal in all non-accelerating frames. The paper also included Lorentz transformations, transformations between reference systems moving at constant velocities relatively to each other, as a necessity to make the speed of light constant in all reference systems.

The theory of Lorentz transformations was completed by Joseph Larmor in 1898[17]. It involves both time dilation and length contraction. Time dilation states that time develops at different rates in frames moving relatively to each other. Imagine taking a trip into space and back. Because you travel with high velocities, you would observe time developing slower than the remainders at earth. Length contraction, on the other hand, states that the length of an object depends on the reference frame of the observer; if an object is moving parallel with its length, an observer would measure the length to be shorter than it would be in the object's rest frame.

The most famous formula of Einstein's paper from 1905 is the mass-energy equivalence, $E = mc^2$, which states that mass is energy and vice versa. This implies that all particles with mass have energy, even at rest, and it is of massive importance in relativistic calculations.

1.3.1 The Merge of Special Relativity and Quantum Mechanics

When Schrödinger failed to merge his own quantum equation with special relativity in 1925, scientists like Oskar Klein, Walter Gordon and Paul Dirac stood ready and took on the challenge. Oskar Klein and Warren Gordon proposed the Klein-Gordon equation in 1926, which was meant to be a relativistic analogue to Schrödinger's equation for a free electron. Their equation yielded both negative energy solutions and negative probabilities, which was believed to be unphysical.

Therefore, Paul Dirac proposed a first order 4-vector equation, the Dirac equation for a free electron, to get rid of these unphysical properties from the Klein-Gordon equation[18]. The Dirac equation did not remove these, but used them to predict the existence of antiparticles. The existence of antiparticles shows the limitations of classical quantum mechanics, because the number of particles no longer is conserved. Conservation of the number of particles is a necessity in classical quantum mechanics, but it is not conserved in physical processes like pair creation and annihilation of particles and its antiparticles. Quantum mechanics fails in describing these processes.

The Dirac equation has furthermore been used to give a description of hydrogenlike atoms, atoms with only one electron, consistent with both quantum mechanics and special relativity. Full solutions and descriptions of hydrogenlike atoms have already been proposed in the field of atomic physics[19]. It is shown that the expected ground state energy, the lowest energy, is lower than the corresponding expected ground state energy in the Schrödinger equation for hydrogenlike atoms with high nucleus charge[19]. This relativistic effect stems from the high velocity of the electron due to the strong Coulomb potential. These observations are consistent with observations of relativistic chemistry[20].

It is also of interest to study how high-energy lasers will interact with hydrogenlike atoms, as it is reasonable to suspect that the high-energy interaction with hydrogenlike atoms may yield relativistic effects. A classical analogue to this interaction may be a free electron interacting with an oscillating homogenous electric field. The electron will quiver in a velocity proportional to the strength of the electric field. A very strong electric field can induce high velocities of the electron, and the electron will need a relativistic description. The quantum description of hydrogenlike atoms interacting with a high-energy laser pulse is a bit more complicated, and it is

expected that the Schrödinger equation is insufficient to describe such a system.

1.4 The Status of the Research Today

The Dirac equation with laser-atom interactions has received increased scientific attention in the field of atomic physics in the latest years, as it now is possible to produce high-energy pulses. It will then be possible to experimentally measure properties of the high-energy interaction between pulses and atoms and compare it with corresponding results from the Dirac equation.

The inclusion of laser-atom interactions in the Dirac equation for hydrogenlike atoms has proven itself to be difficult to solve and approximations have been necessary. The laser field is usually modelled as a classical, continuous wave, which we also do in this thesis. Furthermore, the most common approximation is the dipole approximation, which assumes that the spatial part of the pulse can be neglected. If the spatial part is neglected, then the magnetic field also disappears. The magnetic field is of great importance in high-energy interactions, as it approaches a substantial fraction of the mass energy. This thesis will though present numerical solutions within the dipole approximation, which results in limited precision for truly intense lasers.

The atomic physics group at Stockholm University has managed to solve the Dirac equation with laser-atom interactions beyond the dipole approximation, but the solution was shown to be insufficient for truly strong laser pulses with peak electric field strength above 90 atomic units[21]. It was suggested by the atomic physics group at the university of Bergen to include the laser pulse interaction in the Dirac equation through a gauge transformation, a special form of Lorentz transformations, of the laser pulse[22,23]. This gauge transformation has shown promise as it introduces the possibility to include truly strong laser pulses and perhaps observe more distinct relativistic effects.

This thesis models hydrogenlike atoms exposed to high-energy pulses in the Dirac equation and proposes numerical solutions. The high-energy pulse is modelled as a classical wave and is included in the Dirac equation by using the method of minimal coupling. By applying and implementing the minimal coupling, it is possible to study the properties of the interaction between hydrogenlike atoms and relatively strong laser pulses. The results is expected to yield an indication of how the relativistic effects express itself.

The Dirac equation, without laser-atom interactions, is first solved for stationary states, and the relativistic ground state energy is compared with the non-relativistic ground state energy yielded from the Schrödinger equation. Furthermore, the Dirac equation, with laser-atom interactions, is propagated in time by using the Magnus expansion of lowest order[21, 24] and Krylov subspace methods[25]. Probability of photoionization of the hydrogen ground state in the Dirac equation will be compared with corresponding results from the Schrödinger equation with and without relativistic corrections. Specifically will the laser be modeled with angular frequency $\omega = 50$ and various field strengths.

Chapter 2

Theory

We will in this chapter introduce theory and concepts necessary to grasp the Dirac equation.

2.1 Atomic Units

It is normal to use atomic units in the field of atomic physics, unlike many other scientific fields which use S.I. units. We will use atomic units in this thesis, as it simplifies many expressions later on. The most important conventions in atomic units are given in table 2.1.

Table 2.1: Conventions in atomic units

Symbol	**Description**	**Value in S.I.**	**Value in Atomic units**
$\hbar$	Reduced Planck's constant	$6.626 \cdot 10^{-34} Js$	1 a.u.
m_e	Electron rest mass	$9.109 \cdot 10^{-31} kg$	1 a.u.
k_e	Coloumb's constant	$8.99 \cdot 10^{9} \frac{Nm^2}{C^2}$	1 a.u.
c	Speed of light	$3 \cdot 10^{8} \frac{m}{s}$	137 a.u.
e	Electron charge	$1.602 \cdot 10^{-19} C$	1 a.u.
a_0	Bohr radius	$5.292 \cdot 10^{-11} m$	1 a.u.

The constants which are unity in table 2.1, will from now on be left out from mathematical descriptions.

2.2 Relativistic Notation

Relativistic Notation, often called Einstein's notation, is a notation that compresses summation over arbitrary vectors, often four-vectors. A four-vector is like its name suggest, a vector with four components. It transforms in a specific way under Lorentz transformations, which are coordinate transformations between coordinate frames that have constant velocity relative to one another[19].

The space-time four-vector is defined as

$$\begin{aligned} x^\mu &= (ct, \vec{r}) = (ct, x, y, z) = (ct, r^i) \\ x_\mu &= (ct, -\vec{r}) = (ct, -x, -y, -z) = (ct, -r^i) \end{aligned} \tag{2.1}$$

Where μ ranges from 0 to 3, and i ranges from 1 to 3. Other encountered four-vectors in this thesis are

$$\begin{aligned} energy-momentum: &\quad p^\mu = (E, \vec{p}) \\ four-gradient: &\quad \partial^\mu = \left(\frac{\partial}{\partial t}, -\vec{\nabla}\right) \\ four-vectorpotential: &\quad A^\mu = (A_0, \vec{A}) = (\phi, \vec{A}) \end{aligned}$$

The invariant length of the four-vector x is given by

$$x^2 = x \cdot x = x_\mu x^\mu = (ct)^2 - \vec{r}^2 = (ct)^2 - x^2 - y^2 - z^2 = g_{\mu\nu} x^\mu x^\nu \tag{2.2}$$

Repeated indexes, like μ in equation (2.2), always indicate a summation. $g_{\mu\nu}$ is the (μ, ν) element of the g-matrix;

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

In general one can state that the dot product between two four-vectors $\alpha^\mu = [\alpha_0, \vec{\alpha}]$ and $\beta^\mu = [\beta_0, \vec{\beta}]$

$$\alpha_\mu \beta^\mu = \alpha_0 \beta_0 - \vec{\alpha} \cdot \vec{\beta} = \alpha_0 \beta_0 - \alpha_i \beta_i \tag{2.3}$$

The lest term in (2.3), $\alpha_i \beta_i$, refers to a dot product between two three-component vectors. We emphasize that the repeated index i indicates a summation over $i = 1, 2, 3$. For simplicity, this shorthand notation will be used extensively throughout the thesis.

We will also encounter cross-products between two arbitrary three-vectors $\vec{\alpha} = [\alpha_1, \alpha_2, \alpha_3]$ and $\vec{\beta} = [\beta_1, \beta_2, \beta_3]$. The i-th component of the cross product $\vec{\alpha} \times \vec{\beta}$ can be written as

$$\left(\vec{\alpha} \times \vec{\beta}\right)_i = \epsilon_{ijk}\alpha_j\beta_k, \tag{2.4}$$

where j and k also ranges from 1 to 3. ϵ_{ijk} is the Levi-Cevita symbol[26] and it is anti-symmetric in permutations of i, j and k.

2.3 Postulates of Quantum Mechanics

Relativistic quantum mechanics differs from non-relativistic quantum mechanics in some regards, but they also share many common principles. Four of the postulates from non-relativistic quantum mechanics is valid for relativistic quantum mechanics[27]:

Postulate 1 The state of a physical system is described by a wavefunction $\psi(\vec{r}, t)$, that is dependent on the position $\vec{r}$ of the particle(s) present and on time t. The probability of finding a particle in a volume element dV at position $\vec{r}$ at a time t is given by $\psi(\vec{r}, t)^\dagger \psi(\vec{r}, t) dV$.

The wavefunction $\psi(\vec{r}, t)$ must be normalized, i.e.

$$1 = \int_V dV \psi(\vec{r}, t)^\dagger \psi(\vec{r}, t), \tag{2.5}$$

as the probability to find a particle anywhere in the volume V must be unity.

Postulate 2 To every classical physical observable, there exist a corresponding hermitian operator. See table 2.2 for important physical observables with their corresponding hermitian operator.

Table 2.2: Physical observables with corresponding hermitian operator

Physical Observables	**Operator symbol**	**Hermitian operator**
Momentum	$\hat{p}$	$-i\nabla$
Position	$\hat{r}$	Multiply with r
Kinetic energy	$\hat{T}$	$-\frac{1}{2m}\nabla^2$
Potential energy	$\hat{V}$	Multiply with $V(r)$
Angular momentum	$\hat{L}_i$	$\epsilon_{ijk} r_j \nabla_k$

Postulate 3 Upon measurement of physical observables, the only values that will ever be observed are the eigenvalues λ of the physical observable's corresponding hermitian operator A.

$$\hat{A}\psi = \lambda\psi \tag{2.6}$$

Postulate 4 The average of an observable, corresponding to the hermitian operator $\hat{A}$, is given by

$$< A >= \int_V dV \psi^\dagger \hat{A}\psi \tag{2.7}$$

Postulate 5 The fifth postulate is not valid for relativistic quantum mechanics, although it is of importance. It states that a wavefunction ψ evolves in time according to the time-dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \hat{H}\psi(\vec{r},t) \tag{2.8}$$

where $\hat{H} = \hat{T} + \hat{V}$ is the non-relativistic Hamiltonian. If the Hamiltonian $\hat{H}$ is time-independent, one can search for stationary solutions of the form $\psi(\vec{r},t) = \exp(-iEt)\psi(\vec{r})$, which inserted back into the time-dependent Schrödinger equation (2.8) yields the time-independent Schrödinger equation.

$$\hat{H}\psi(\vec{r}) = E\psi(\vec{r}) \tag{2.9}$$

As the energy E changes form in the relativistic description, one needs to find a new Hamiltonian $\hat{H_D}$ which is valid in the relativistic regime. We assume that this new relativistic Hamiltonian $\hat{H_D}$ must satisfy the same time development as the non-relativistic Hamiltonian $\hat{H}$;

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \hat{H}_D\psi(\vec{r},t) \tag{2.10}$$

This equation, (2.10), will be derived later on and will be called the time-dependent Dirac equation.

It is also worth mentioning if two hermitian operators $\hat{A}$ and $\hat{B}$ commute, that is

$$\left[\hat{A},\hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} = 0, \tag{2.11}$$

they can have simultaneous eigenvalues for ψ.

2.4 From the Schrödinger equation to the Dirac Equation

We stated earlier that we need to find a Hamiltonian which is valid in the relativistic regime. We start from the non-relativistic description of the kinetic energy of a free electron:

$$E = \frac{p^2}{2} \tag{2.12}$$

and impose $E \rightarrow i\frac{\partial}{\partial t}$ and $p \rightarrow -i\nabla$. We let both sides act on a wavefunction $\psi(\vec{r},t)$ and it yields the time-dependent Schrödinger equation.

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{1}{2}\nabla^2\psi(\vec{r},t) \tag{2.13}$$

The energy relation in (2.12) is not valid in the relativistic regime, as mass also is considered as energy (here $m_e = 1$ due to atomic units). The correct relativistic energy E is given by the energy-momentum relation[18]

$$E^2 = p^2c^2 + c^4 \tag{2.14}$$

We let $E \rightarrow i\frac{\partial}{\partial t}$ and $p \rightarrow -i\nabla$ as we did for the time-dependent Schrödinger equation (2.13). Thereafter we let it act on a wavefunction $\psi(\vec{r},t)$, and it yields the Klein-Gordon equation,

$$-\frac{\partial^2}{\partial t^2}\psi(\vec{r},t) = \left(-\nabla^2c^2 + c^4\right)\psi(\vec{r},t) \tag{2.15}$$

The Klein Gordon equation (2.15) was historically discarded, as it yielded negative probability densities. It is shown in the following paragraphs.

If the probability ρ of finding a particle in a finite volume V increases, it must be equal to the probability current $\vec{j}$ through the surface F of the volume V.

$$\frac{\partial}{\partial t}\int_V \rho\, d^3r = -\int_F \vec{j}\cdot d\vec{f} \tag{2.16}$$

The divergence theorem states $\int_V \nabla\cdot\vec{j}\,d^3r = \int_F \vec{j}\cdot d\vec{f}$, which transforms equation (2.16) to

$$\int_V \left(\frac{\partial}{\partial t}\rho + \nabla\cdot\vec{j}\right) d^3r = 0 \tag{2.17}$$

As this should be valid for any volume V, the integrand must be zero:

$$\frac{\partial}{\partial t}\rho + \nabla\cdot\vec{j} = 0 \tag{2.18}$$

This is known as the continuity equation. It would be advantageous to retrieve this equation from the Klein Gordon equation, (2.15), as the probability density ρ can be analyzed. Taking ψ^* times (2.15), ψ times the complex conjugate of (2.15), and subtracting these, yields:

$$\frac{\partial}{\partial t}\left(\frac{i}{2c^2}\left(\psi^*\frac{\partial}{\partial t}\psi-\psi\frac{\partial}{\partial t}\psi^*\right)\right)+\nabla\cdot\left(\frac{1}{2i}\left(\psi^*\nabla\psi-\psi\nabla\psi^*\right)\right)=0 \tag{2.19}$$

We would like to interpret $\frac{i}{2c^2}\left(\Psi^*\frac{\partial}{\partial t}\Psi-\Psi\frac{\partial}{\partial t}\Psi^*\right)$ as probability density ρ according to equation (2.18). It is not positive definite, which was historically believed to be unphysical and the Klein Gordon equation was therefore discarded. The Klein-Gordon is to this day interpreted as an equation describing spin-0 particles[28]. It was believed that the quadratic partial derivatives in the Klein Gordon equation (2.15) was the reason for these negative probability densities. Dirac therefore proposed a first order partial differential equation, in order to get rid of the quadratic partial derivatives in the Klein Gordon equation.

Dirac proposed an equation of the form[19]

$$i\frac{\partial}{\partial t}\psi(r,t)=-ic\left(\sum_{i=1}^{N}\alpha_i\frac{\partial}{\partial x_i}\psi(r,t)\right)+\beta c^2\psi(r,t), \tag{2.20}$$

where the coefficients α_i and β can not just be constants, as the equation would not be invariant under spatial rotations. Therefore was α_i and β proposed to be constant NxN matrices, which correspondingly forces ψ to be a column vector with N components,

$$\psi=\begin{bmatrix}\psi_1\\ \psi_2\\ \vdots\\ \psi_N\end{bmatrix} \tag{2.21}$$

We can rewrite equation (2.20) as

$$\frac{\partial}{\partial t}\psi+c\left(\vec{\alpha}\cdot\nabla\right)\psi+ic^2\beta\psi=0, \tag{2.22}$$

where $\vec{\alpha}=\left(\alpha_1,\alpha_2,\alpha_3\right)$. Each component of ψ, ψ_i, should satisfy the Klein Gordon equation (2.15), as the energy relation $E^2=c^2p^2+c^4m^2$ must be satisfied. Furthermore, we search for a positive definite probability density ρ, to avoid the downfall of the Klein-Gordon equation (2.15).

$$\rho=\psi^\dagger\psi=\sum_{i=1}^{N}\psi_i^*\psi_i=\sum_{i=1}^{N}|\psi_i|^2>0 \tag{2.23}$$

We assume that the probability current $\vec{j}$ has the form

$$\vec{j} = c\psi^\dagger \vec{\alpha}\psi \tag{2.24}$$

The probability density ρ and the probability current $\vec{j}$ must satisfy the continuity equation (2.18). We show this by taking the hermitian conjugate of equation (2.22)

$$\frac{\partial}{\partial t}\psi^\dagger + c\nabla\psi^\dagger \cdot \vec{\alpha}^\dagger - ic^2\psi^\dagger\beta^\dagger = 0 \tag{2.25}$$

and use the same method as we did for the Klein-Gordon equation. We multiply equation (2.22) with $\psi^\dagger$ from the left and equation (2.25) with ψ from the right.

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) + c\left(\psi^\dagger(\vec{\alpha}\cdot\vec{\nabla})\psi + \nabla\psi^\dagger\cdot\vec{\alpha}^\dagger\psi\right) + ic^2\left(\psi^\dagger\beta\psi - \psi^\dagger\beta^\dagger\psi\right) = 0 \tag{2.26}$$

We impose $\vec{\alpha} = \vec{\alpha}^\dagger$ and $\beta = \beta^\dagger$, which yields

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) + \nabla\cdot\left(c\psi^\dagger\vec{\alpha}\psi\right) = 0, \tag{2.27}$$

which shows that equation (2.23) and (2.24) combined satisfy the continuity equation, and we do indeed have a positive definite probability density.

We have already put restrictions on $\vec{\alpha}$ and β, as $\vec{\alpha} = \vec{\alpha}^\dagger$ and $\beta = \beta^\dagger$. It was mentioned earlier that each component of ψ needs to satisfy the Klein Gordon (2.15). In order for ψ to do this, $\vec{\alpha}$ and β still need to be further determined. We start by transforming equation (2.22) to the Klein-Gordon equation by acting on it with "$\frac{\partial}{\partial t} - c\alpha\cdot\nabla - ic^2\beta$" from the left;

$$\frac{\partial^2}{\partial t^2}\psi = c^2\sum_{j,k}\frac{1}{2}\left(\alpha_j\alpha_k + \alpha_k\alpha_j\right)\frac{\partial^2}{\partial x_j\partial x_k}\psi - \beta^2\psi + ic^4\sum_j\left(\alpha_j\beta + \beta\alpha_j\right)\frac{\partial}{\partial x_j}\psi = 0 \tag{2.28}$$

The right side of this equation should be equal to $\left(c^2\nabla^2 - c^4\right)$ from the Klein-Gordon equation (2.15), so $\vec{\alpha}$ and β must satisfy

$$\frac{1}{2}\left(\alpha_j\alpha_k + \alpha_k\alpha_j\right) = \delta_{j,k}\mathbb{1}_N \tag{2.29}$$

$$\alpha_j\beta + \beta\alpha_j = 0 \tag{2.30}$$

$$\beta^2 = \mathbb{1}_N \tag{2.31}$$

Equation (2.29) implies $\alpha_i^2 = \mathbb{1}_N$, and therefore α and β, according to equation 2.31, must have

eigenvalues equal to ± 1. Equation (2.30) can be rewritten as:

$$\alpha_i \beta = -\mathbb{1}_N \beta \alpha_i \tag{2.32}$$

Taking the determinant of both sides yields

$$\det\left(\alpha_i \beta\right) = \det\left(-\mathbb{1}_N \beta \alpha_i\right) = \left(-1\right)^N \det\left(\beta \alpha_i\right) \tag{2.33}$$

As $\det\left(AB\right) = \det\left(A\right)\det\left(B\right)$ for square matrices A and B of equal size, $\left(-1\right)^N = 1$ according to equation (2.33). Thus N must be even and it can also be shown that we must have $N \geq 4$[19]. Dirac worked out the α and β matrices for $N = 4$ by using the Pauli matrices,

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix}, \quad \text{where } \sigma_i \text{ is a Pauli matrix;} \tag{2.34}$$

$$\begin{aligned} \sigma_1 = \sigma_x &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \sigma_2 = \sigma_y &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\ \sigma_1 = \sigma_x &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned} \tag{2.35}$$

and the β matrix is given by

$$\beta = \begin{bmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{bmatrix} \tag{2.36}$$

Dirac's original equation was given by equation (2.22)

$$\frac{\partial}{\partial t}\psi + c(\vec{\alpha} \cdot \nabla)\psi + ic^2 \beta \psi = 0, \tag{2.37}$$

which is equivalent to

$$\frac{\partial}{\partial t}\psi + c\alpha_i \nabla_i \psi + ic^2 \beta \psi = 0, \tag{2.38}$$

This can be rewritten as

$$i\frac{\partial}{\partial t}\psi = \Big(-ic\alpha_i \nabla_i + c^2 \beta\Big)\psi, \tag{2.39}$$

which is of the same form as the non-relativistic time-dependent Schrödinger equation. We therefore interpret this as the time-dependent Dirac equation. This leads to the assumption of

letting the right side of equation (2.58) be the relativistic Hamiltonian H_D for a free particle.

$$\begin{aligned} H_D &= -ic\alpha_i \nabla_i + c^2\beta \\ &= c\alpha_i \hat{p}_i + c^2\beta \end{aligned} \tag{2.40}$$

We note that equation (2.39) is for a free electron. It has been solved in[19], where it is shown that there actually exist negative energy solutions for a free electron. Dirac introduced the Dirac sea to explain these negative energy solutions. He postulated there exists a 'sea' filled by electrons with negative energies $E < -c^2$. If an electron in the Dirac is exited by a photon to a state with $E > c^2$, as this is an electrons rest energy, it creates a hole in the Dirac sea. After the discovery of the positron in 1932[29] this hole was interpreted as the electrons antiparticle, the positron. The excitation of an electron in the Dirac sea to a state with $E > c^2$ was as a consequence interpreted as a pair-creation process. Analogously was the de-excitaiton of an electron with $E > c^2$ to the the Dirac sea interpreted as an annihilation process. The discovery of positrons lead to the fact that electrons in the Dirac sea could be interpreted as positrons.

As we want to model a hydrogenic atom, we need to include the interaction between the nucleus and the electron into the time-dependent Dirac equation (2.39). We will in this thesis treat the nucleus as a rigid point charge at the origin, so we can interpret the interaction between the nucleus and the electron as a Coulomb interaction. The Coulomb potential $V(r)$ created by the nucleus is given by

$$V(r) = \frac{Z}{|r|}, \tag{2.41}$$

where Z is the number of protons in the nucleus.

We include the Coloumb potential $V(r)$ by the method of minimal coupling,

$$\partial_\mu \to \partial_\mu - iA_\mu \tag{2.42}$$

The method of minimal coupling can be derived from quantum field theory, which we will take a closer look at in the next section.

2.5 Quantum Field Theory and the Method of Minimal Coupling

Quantum Field Theory works under the mathematical description of elementary particles, like electrons, as quantum fields. Every physical particle and wave is interpreted as an excitation of these fields. All quantum systems are described by a Lagrangian, a function that contains all

information about the system[30].

The Lagrangian is defined as the difference between the kinetic- and potential energy;

$$L = T - V, \tag{2.43}$$

where T is the kinetic energy and V is the potential energy. The integral of the Lagrangian over time τ is defined as the action S

$$S = \int_{\tau} L dt \tag{2.44}$$

Imagine a ball being thrown in a gravitational field. The x-coordinate of the ball is described by a function x(t). The trajectory the ball follows, is described by Hamilton's principle of least action. The Hamilton's principle of least action states that the trajectory of a classical particle is chosen such that the action is stationary,

$$\frac{\delta S}{\delta x(t)} = 0 \tag{2.45}$$

When going to 3-dimensional problems in quantum field theory, it is quite useful to work with the Lagrangian density instead of the Lagrangian. This stems from the fact that particles in quantum mechanics are described by probability densities. Let us define a Lagrangian density $\mathscr{L}$ in an arbitrary volume Ω[30],

$$L = \int_{\Omega} d^3 x \mathscr{L}, \tag{2.46}$$

The lagrangian density is dependent on the quantum fields ϕ_r describing the present elementary particles,

$$\mathscr{L} = \mathscr{L}\big(\phi_r, \phi_{r,\mu}\big), \tag{2.47}$$

where $\phi_{r,\mu} = \partial_\mu \phi_r$ and r labels the component of the quantum field ϕ_r. The action integral $S(\Omega)$ in equation (2.44) for an arbitrary volume Ω, expressed by the Lagrangian density, takes the form

$$S(\Omega) = \int_{\Omega} d^4 x \mathscr{L}\big(\phi_r, \phi_{r,\mu}\big) \tag{2.48}$$

The equations of motion, i.e the classical motion equations, can be obtained from applying the variational principle. We vary the field ϕ_r by stating

$$\phi_r(x) \rightarrow \phi_r(x) + \delta\phi_r(x) \tag{2.49}$$

The variation $\delta\phi_r$ should not change the boundary conditions of the field ϕ_r and therefore must

$\delta\phi_r$ be zero at the surface Γ of the region Ω

$$\delta\phi_r(x) = 0, \quad \text{on } \Gamma(\Omega) \tag{2.50}$$

These variations of the field ϕ_r will lead to a variation in the action integral, δS, in equation (2.48). Hamilton's principle states that nature always aligns itself so that this variation δS is zero, i.e. the action integral is stationary,

$$\delta S(\Omega) = 0 \tag{2.51}$$

The variation in δS can be defined as[30]

$$\begin{aligned} \delta S(\Omega) &= \int_\Omega d^4x \Big(\frac{\partial \mathscr{L}}{\partial \phi_r} \delta\phi_r + \frac{\partial \mathscr{L}}{\partial \phi_{r,\alpha}} \delta\phi_{r,\alpha} \Big) \\ &= \int_\Omega d^4x \Big(\frac{\partial \mathscr{L}}{\partial \phi_r} - \frac{\partial}{\partial x^\alpha} \Big(\frac{\partial \mathscr{L}}{\partial \phi_{r,\alpha}} \Big) \Big) \delta\phi_r + \int_\Omega d^4x \frac{\partial}{\partial x^\alpha} \Big(\frac{\partial \mathscr{L}}{\partial \phi_{r,\alpha}} \delta\phi_r \Big) \end{aligned} \tag{2.52}$$

The last line is obtained by the method of partial integration. The last term vanishes as a consequence of Gauss' divergence theorem and equation (2.50). As $\delta S(\Omega)$ should vanish for any region Ω, one must demand that the integrand is zero:

$$\frac{\partial \mathscr{L}}{\partial \phi_r} - \frac{\partial}{\partial x^\alpha} \Big(\frac{\partial \mathscr{L}}{\partial \phi_{r,\alpha}} \Big) = 0, \quad \text{for } r = 1, ..., N \tag{2.53}$$

This is known as the Euler-Lagrange equations and lead to the equations of motion.

The Lagrangian for a Dirac field describing electrons, is given by[30]

$$\mathscr{L}_D = \bar{\psi} \big(i c \gamma^\mu \partial_\mu - c^2 \big) \psi, \quad \text{where} \quad \bar{\psi} \equiv \psi^\dagger \gamma^0 \tag{2.54}$$

The γ-matrices are defined by the α-matrices and the β-matrix: $\gamma^\mu = \big[\gamma^0, \beta\alpha_i\big]$, where $\gamma^0 = \beta$. One can retrieve the Dirac equation by applying the Euler-Lagrange equations (2.53).

Transformations that leave the Lagrangian $\mathscr{L}_D$ invariant, are called gauge transformations. Physical transformations of the field ψ, like a Coulomb interaction or a classical pulse, are gauge transformations. We propose a local gauge transformation of the field of the form $\psi \rightarrow \psi' = \exp(i\chi(x^\mu))\psi$, which should leave the Lagrangian $\mathscr{L}_D$ invariant, as it is equivalent to rotation of the complex plane on which ψ is defined. It should not induce a difference in physical measurements. The transformation $\psi \rightarrow \psi' = \exp(-i\chi(x^\mu))\psi$ changes the Lagrangian $\mathscr{L}_D$;

$$\begin{aligned} \mathscr{L}_D &= \exp\big(i\chi(x^\mu)\big) \bar{\psi} \Big(i c \gamma^\mu \partial_\mu - c^2 \Big) \psi \exp\big(-i\chi(x^\mu)\big) \\ &= \bar{\psi} \Big(i c \gamma^\mu \big(\partial_\mu - \partial_\mu \chi(x^\mu) \big) - c^2 \Big) \psi \end{aligned} \tag{2.55}$$

The Lagrangian $\mathscr{L}_D$ is invariant by imposing

$$\partial_\mu \to \partial_\mu - iA_\mu \tag{2.56}$$

$$A_\mu \to A_\mu + \partial_\mu \chi(x^\mu), \tag{2.57}$$

where is A_μ a four-vector potential. As $p^\mu = -i\partial_\mu$, it will transform as $p^\mu \to p^\mu + A_\mu$ under the transformation in equation (2.56). This is known as minimal coupling.

2.6 Derivation of the Radial Dirac Equation

In section 2.4 we derived the time-dependent Dirac equation

$$i\frac{\partial}{\partial t}\psi = \Big(-cia_i\nabla_i + c^2\beta\Big)\psi \tag{2.58}$$

The next task is to include a Coulomb potential, so the Hamiltonian has the possibility to describe a hydrogenic atom. This is done through the method of minimal coupling, where $\partial_u \to \partial_\mu - iA_\mu$ in equation (2.58). The Coulomb potential $V(r)$ is a scalar potential, so we state $A = [V,0,0,0]$. The method of minimal coupling from (2.56) therefore takes the form

$$\frac{\partial}{\partial t} \to \frac{\partial}{\partial t} - iV, \tag{2.59}$$

which transforms (2.58) to

$$i\frac{\partial}{\partial t}\psi = \Big(-ic\nabla_i\partial_i - V + c^2\beta\Big)\psi, \tag{2.60}$$

where the right side is the Dirac Hamiltonian H_D for a hydrogenic atom

$$H_D = -icα^i\nabla_i - V + c^2\beta = -ic\vec{\alpha}\cdot\vec{\nabla} - V + c^2\beta \tag{2.61}$$

As the Hamiltonian H_D is time independent, we search for stationary solutions

$$\psi(\vec{r},t) = \exp\big(-iEt\big)\psi(\vec{r}), \tag{2.62}$$

such that

$$H_D\psi(\vec{r}) = E\psi(\vec{r}), \tag{2.63}$$

which is the time-independent Dirac equation, and we seek to solve it. The Dirac Hamiltonian H_D is at its current form too difficult to solve, and we therefore seek to simplify it.

It is natural to start to search for observables that are constants of motion, which is the same

as searching for hermitian operators that commute with the Hamiltonian H_D. If an operator A commute with the Hamiltonian H_D, the operator A and the Hamiltonian H_D can have simultaneously eigenvalues. If we find such operators A, we can use its eigenvalues to fix the state ψ.

It is natural to start with the operators that fix the hydrogenic wavefunction ψ^{NR}, where NR is an abbreviation for non-relativistic, in the NR Schrödinger equation, namely the total angular momentum $\vec{J}$, the angular momentum $\vec{L}$ and the spin $\vec{S}$. It has to be decided whether these operators are constants of motion in the Dirac equation or not.

The angular momentum is defined as it is in classical mechanics;

$$\vec{L} = \vec{r} \times \vec{p} = -i\vec{r} \times \vec{\nabla}, \tag{2.64}$$

where $\vec{L} = \left(L_x, L_y, L_z\right)$. L_z is used to fix the non-relativistic state ψ^{NR} in the non-relativistic, hydrogenic Schrödinger equation;

$$L_z \psi^{NR}_{m_l} = m_l \psi^{NR}_{m_l}, \tag{2.65}$$

where m_l is the magnetic quantum number.

It is shown in appendix A.1.1 that the commutator for H and $\vec{L}$ is

$$\left[H_D, \vec{L}\right] = -c\vec{\alpha} \times \vec{\nabla} \neq 0, \tag{2.66}$$

which means that $\vec{L}$ is not a constant of motion and neither of its components can fix the relativistic state ψ, like it does for ψ^{NR} in equation (2.65). $\vec{L}^2 = \vec{L} \cdot \vec{L} = L_x^2 + L_y^2 + L_z^2$ is also used to fix the non-relativistic state ψ^{NR}_{l,m_l} in the non-relativistic Schrödinger equation;

$$\vec{L}^2 \psi^{NR}_{l,ml} = l(l+1)\psi^{NR}_{l,m_l}, \tag{2.67}$$

where l is the orbital angular momentum quantum number. It does not, however, commute with the relativistic Hamiltonian H_D (see appendix A.1.2 for proof) and is therefore not adequate as a quantum number for the relativistic state ψ.

$$\left[H_D, \vec{L}^2\right] = 2i\Big(\left(\vec{r} \cdot \alpha\right)\left(\vec{\nabla} \cdot \vec{\nabla}\right) - \left(\vec{r} \cdot \vec{\nabla}\right)\left(\vec{\alpha} \cdot \vec{\alpha}\right)\Big) \neq 0 \tag{2.68}$$

As $\vec{L}$ and $\vec{L}^2$ has shown themselves to be inadequate to fix the relativistic state ψ, it is time to

investigate the spin $\vec{S}$. The spin $\vec{S}$ operator is defined as;

$$\vec{S} = \frac{1}{2}\vec{\sigma}, \tag{2.69}$$

where $\vec{S} = \left(S_x, S_y, S_z\right)$ and $\vec{\sigma} = \left[\sigma_x, \sigma_y, \sigma_z\right]$ are $4x4$ Pauli matrices. S_z is used in non-relativistic quantum mechanics to fix the state ψ^{NR};

$$S_z \psi^{NR}_{l,m_l,m_s} = \hbar m_s \psi^{NR}_{l,m_l,m_s}, \tag{2.70}$$

where m_s is the electron spin quantum number. In non-relativistic quantum mechanics, an electron is either spin up $\left(m_s = \frac{1}{2}\right)$ or spin down $\left(m_s = -\frac{1}{2}\right)$. This does not apply in relativistic quantum mechanics as $\vec{S}$ is not a constant of motion. $\vec{S}$ does not commute with H_D. The derivation is shown in appendix A.1.3

$$\left[H_D, \vec{S}\right] = c\vec{\alpha} \times \vec{\nabla} \neq 0 \tag{2.71}$$

It is known in non-relativistic quantum mechanics that $\vec{S}^2$ yields $s = \frac{1}{2}$, when acting on a non-relativistic state ψ^{NR},

$$\vec{S}^2 \psi^{NR} = s(s+1)\psi^{NR}, \tag{2.72}$$

where $s = \frac{1}{2}$ is the electron spin. This is also true for a relativistic electron as $\vec{S}^2$ is constant of motion for the Dirac Hamiltonian H_D. See appendix A.1.4 for proof.

$$\left[H_D, \vec{S}^2\right] = 0 \tag{2.73}$$

The last operator from non-relativistic quantum mechanics that can fix the state ψ is the total angular momentum operator $\vec{J}$. The total angular momentum operator $\vec{J}$ is defined as the sum of angular momentum $\vec{L}$ and the spin $\vec{S}$;

$$\vec{J} = \vec{S} + \vec{L} \tag{2.74}$$

$\vec{J}$ is obviously constant of motion as it commutes with H_D;

$$\begin{aligned} \left[H_D, \vec{J}\right] &\overset{(2.74)}{=} \left[H_D, \vec{L} + \vec{S}\right] \\ &= \left[H_D, \vec{L}\right] + \left[H_D, \vec{S}\right] \\ &\overset{(2.71),(2.66)}{=} c\vec{\alpha} \times \vec{\nabla} + c\vec{\alpha} \times \vec{\nabla} \\ &= 0 \end{aligned} \tag{2.75}$$

It is therefore possible to fix the state ψ by using the total angular momentum $\vec{J}$ in two different

ways, by the quantum numbers j and m_j;

$$\vec{J}^2\psi = j(j+1)\psi_{jm_j} \tag{2.76}$$

$$J_z\psi = m_j\psi_{jm_j}, \tag{2.77}$$

which is equivalent to the definition in non-relativistic quantum mechanics. The quantum numbers j and m_j are not enough to fix the state ψ_{j,m_j} in non-relativistic quantum mechanics as the angular momentum l can be parallel and anti-parallel to the spin $s = \frac{1}{2}$

$$j = l \pm \frac{1}{2} \tag{2.78}$$

This calls for a new operator in relativistic quantum mechanics that analogously can differ between states where l and s are parallel or anti-parallel. Such an operator is the Runge-Lenz vector K_R[31];

$$K_R = \beta\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\beta \tag{2.79}$$

It would be advantageous to make use of this operator in the Dirac Hamiltonian H_D, as one can show that it is a constant of motion (see appendix A.1.5). Let us assume that K_R acting on a relativistic state ψ yields a quantum number κ_R;

$$K_R\psi_{\kappa_R} = -\kappa_R\hbar\psi_{\kappa_R} \tag{2.80}$$

We can denote a state by its fixed quantum numbers j, m_j and κ_R;

$$\psi = \psi_{\kappa_R,j,m_j} \tag{2.81}$$

There exist a relation between the quantum number κ_R and the total angular momentum j. Let us square the K_R operator in equation (2.79);

$$\begin{aligned} K_R^2\psi_{\kappa_R,j,m_j} &= \left(\beta\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\beta\right)^2\psi_{\kappa_R} \\ &= \kappa_R^2\psi_{\kappa_R,j,m_j} \end{aligned} \tag{2.82}$$

Equation (2.82) can also be written as; (See appendix A.1.6 for derivation)[32]

$$K_R^2\psi_{\kappa_R,j,m_j} = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + 1\right)\psi_{\kappa_R,j,m_j} = \kappa_R^2\psi_{\kappa_R,j,m_j} \tag{2.83}$$

We reformulate (2.76) to a similar form, (See appendix A.1.7 for derivation)

$$\vec{J}^2\psi_{\kappa_R,j,m_j} = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + \frac{3}{4}\right)\psi_{\kappa_R,j,m_j} = j(j+1)\psi_{\kappa_R,j,m_j}, \tag{2.84}$$

which gives the relation

$$\begin{aligned} \kappa_R^2 &= j(j+1) + \frac{1}{4} \\ &\Leftrightarrow \\ \kappa_R &= \pm\left(j + \frac{1}{2}\right) \end{aligned} \tag{2.85}$$

which states that the quantum number κ_R has integer values unequal to zero; $\kappa_R = \pm1, \pm2, \pm3, ...$ It is advantageous to define a new operator K, which is given by

$$K = -\beta K_R = -\left(\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\right) \tag{2.86}$$

K^2 is equal to K_R^2 and hence has eigenvalues κ^2 equal κ_R^2:

$$K^2\psi_{\kappa_R,j,m_j} = \kappa^2\psi_{\kappa_R,j,m_j}, \tag{2.87}$$

where

$$\kappa^2 = \kappa_R^2, \tag{2.88}$$

which according to (2.85) yields

$$\kappa = \pm\left(j + \frac{1}{2}\right) \tag{2.89}$$

The next natural step is to figure out the eigenvalues of K acting on ψ_{κ_R,j,m_j}. We know from earlier derivations that ψ_{κ_R,j,m_j} is a vector with four components. We split it into two two-component parts;

$$\psi_{\kappa_R,j,m_j} = \begin{pmatrix} \psi_{A_{\kappa_R,j,m_j}} \\ \psi_{B_{\kappa_R,j,m_j}} \end{pmatrix} \tag{2.90}$$

We established the relation between the K-operator and the Runge-Lenz vector K_R in (2.86);

$$\begin{aligned} K\psi_{\kappa_R,j,m_j} &= -\beta K_R\psi_{\kappa_R,j,m_j} \\ &= -\begin{pmatrix} \mathbb{1}_{2\times2} & 0 \\ 0 & -\mathbb{1}_{2\times2} \end{pmatrix}(-\kappa_R)\begin{pmatrix} \psi_{A_{\kappa_R,j,m_j}} \\ \psi_{B_{\kappa_R,j,m_j}} \end{pmatrix} \\ &= \begin{pmatrix} \kappa_R\psi_{A_{\kappa_R,j,m_j}} \\ -\kappa_R\psi_{B_{\kappa_R,j,m_j}} \end{pmatrix} \end{aligned} \tag{2.91}$$

It is clear that the K-operator does not yield mutual eigenvalues for ψ_A and ψ_B. $\psi_{A_{\kappa_R,j,m_j}}$ and $\psi_{B_{\kappa_R,j,m_j}}$ are though separate eigenstates of K.[32].

$$\begin{aligned} K\psi_{A_{\kappa_R,j,m_j}} &= \kappa_R \psi_{A_{\kappa_R,j,m_j}} \\ K\psi_{B_{\kappa_R,j,m_j}} &= -\kappa_R \psi_{B_{\kappa_R,j,m_j}} \end{aligned} \tag{2.92}$$

As both components are separate eigenstates of the K-operator, one can propose

$$K\psi_{A\,/\,B,\kappa,j,m} = \kappa\psi_{A\,/\,B,\kappa,j,m}, \tag{2.93}$$

where κ has replaced κ_R as quantum number fixing the state. From (2.92) it is evident that the state ψ must take the form.

$$\psi = \begin{pmatrix} \psi_{A_{\kappa,j,m_j}} \\ \psi_{B_{-\kappa,j,m_j}} \end{pmatrix} \tag{2.94}$$

We can rewrite the definition of the K-operator from equation (2.86), (See appendix A.1.10 for derivation).

$$K = -\left(\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\right) = -\left(\vec{\sigma}\cdot\vec{L} + 1\right) \tag{2.95}$$

$\vec{\sigma}\cdot\vec{L}$ is scalar that can determine whether the spin $\vec{S}$ and $\vec{L}$ are parallel or anti-parallel. Hence the K-operator can differ between states having parallel or anti-parallel spin $\vec{S}$ and angular momentum $\vec{L}$. As ψ_A and ψ_B have opposite sign of κ, it indicates that one of the states have parallel spin $\vec{S}$ and angular momentum $\vec{L}$ and the other anti-parallel. Equation (2.78) stated that $j = l \pm \frac{1}{2}$ in non relativistic quantum mechanics. It stands to reason to investigate whether it is valid in relativistic quantum mechanics as well.

We start by acting with $\vec{L}^2$ on ψ[32];

$$\begin{aligned} \vec{L}^2 \begin{pmatrix} \psi_{A_{\kappa,j,m_j}} \\ \psi_{B_{-\kappa,j,m_j}} \end{pmatrix} &= \left(\vec{J}^2 - \vec{\sigma}\cdot\vec{L} - \frac{\sigma^2}{4}\right) \begin{pmatrix} \psi_{A_{\kappa,j,m_j}} \\ \psi_{B_{-\kappa,j,m_j}} \end{pmatrix} \\ &= \left(j\left(j+1\right) - \vec{\sigma}\cdot\vec{L} - \frac{3}{4}\right) \begin{pmatrix} \psi_{A_{\kappa,j,m_j}} \\ \psi_{B_{-\kappa,j,m_j}} \end{pmatrix} \\ &\overset{(2.95)}{=} \left(j\left(j+1\right) + K + \frac{1}{4}\right) \begin{pmatrix} \psi_{A_{\kappa,j,m_j}} \\ \psi_{B_{-\kappa,j,m_j}} \end{pmatrix} \\ &= \begin{pmatrix} \left(j\left(j+1\right) + \kappa + \frac{1}{4}\right)\psi_{A_{\kappa,j,m_j}} \\ \left(j\left(j+1\right) - \kappa + \frac{1}{4}\right)\psi_{B_{-\kappa,j,m_j}} \end{pmatrix} \end{aligned} \tag{2.96}$$

$\kappa = \pm\left(j+\frac{1}{2}\right)$ according to (2.89). These relations inserted back into (2.96) yields

$$\begin{aligned}\vec{L}^2\begin{pmatrix}\psi_{A_{\kappa,j,m_j}}\\ \psi_{B_{-\kappa,j,m_j}}\end{pmatrix} &= \begin{pmatrix}\left(j(j+1)\pm\left(j+\frac{1}{2}\right)+\frac{1}{4}\right)\psi_{A_{\kappa,j,m_j}}\\ \left(j(j+1)\mp\left(j+\frac{1}{2}\right)+\frac{1}{4}\right)\psi_{B_{-\kappa,j,m_j}}\end{pmatrix}\\ &= \begin{pmatrix}\left(j^2+j\pm j\pm\frac{1}{2}+\frac{1}{4}\right)\psi_{A_{\kappa,j,m_j}}\\ \left(j^2+j\mp j\mp\frac{1}{2}+\frac{1}{4}\right)\psi_{B_{-\kappa,j,m_j}}\end{pmatrix}\end{aligned} \tag{2.97}$$

The relation $l = j\pm\frac{1}{2}$ is known from non-relativistic quantum mechanics. It is verified that it still holds in relativistic quantum mechanics as

$$l(l+1) = \left(j\pm\frac{1}{2}\right)\left(j\pm\frac{1}{2}+1\right) = j^2+j\pm j\pm\frac{1}{2}+\frac{1}{4} \tag{2.98}$$

These results leads to the fact that both ψ_A and ψ_B are eigenstates of $\vec{L}^2$, but with correspondingly different eigenvalues l_A and l_B;

$$\begin{aligned}\vec{L}^2\psi_{A_{\kappa,j,m_j}} &= l_A(l_A+1)\psi_{A_{\kappa,j,m_j}}\\ \vec{L}^2\psi_{B_{-\kappa,j,m_j}} &= l_B(l_B+1)\psi_{B_{-\kappa,j,m_j}}\end{aligned} \tag{2.99}$$

This gives reason to let the angular part of $\psi_{A_{\kappa,j,m_j}}$ be a linear combination of spherical harmonics with $l = l_A$. The angular part of $\psi_{B_{-\kappa,j,m_j}}$ can expressed as a linear combination of spherical harmonics with $l = l_B$. The states are not fixed for m_s and m_l, but we do know that $m_s+m_l = m_j$. Let us denote the angular part of $\psi_{A_{\kappa,j,m_j}}$ as X_{κ,j,m_k} and $\psi_{B_{-\kappa,j,m_j}}$ as $X_{-\kappa,j,m_j}$

$$\begin{aligned}X_{\kappa,j,m_j} &= \sum_{m_s,m_l}\langle l_\kappa, m_{l_\kappa}, s=\frac{1}{2}, m_s|j,m_j\rangle Y_{l_\kappa,m_{l_\kappa}}\chi_{m_s}\\ X_{-\kappa,j,m_j} &= \sum_{m_s,m_l}\langle l_{-\kappa}, m_{l_{-\kappa}}, s=\frac{1}{2}, m_s|j,m_j\rangle Y_{l_{-\kappa},m_{l_{-\kappa}}}\chi_{m_s},\end{aligned} \tag{2.100}$$

where $\langle l, m_l, s=\frac{1}{2}, m_s|j,m_j\rangle$ are Clebsch-Gordan coefficients and χ_{m_s} is the part describing the electron's spin $\vec{S}$ projected along the z-axis.

It is useful to find the relation between the quantum number κ and the angular momentum quantum number l for the upper and lower component. As $KX_{\kappa,j,m_j} = \kappa X_{\kappa,j,m_j}$, where $K = -\vec{\sigma}\cdot\vec{L}-1$, one can determine whether the spin $\vec{S}$ and the angular momentum $\vec{L}$ are parallel or anti-parallel. $\kappa > 0$ must state that $j = l-\frac{1}{2}$ and $\kappa < 0$ must state that $j = l+\frac{1}{2}$. It becomes

evident, as $\kappa = \pm(j+\frac{1}{2})$, that

$$\kappa = \begin{cases} l, & j = l - \frac{1}{2} \\ -\big(l+1\big), & j = l + \frac{1}{2} \end{cases}$$

We can now assume that our relativistic state ψ_{n,κ,j,m_j} has the form

$$\psi_{n,\kappa,j,m_j}(\vec{r}) = \frac{1}{r}\begin{pmatrix} P_{n,\kappa}(r)X_{\kappa,j,m_j} \\ iQ_{n,\kappa}(r)X_{-\kappa,j,m_j} \end{pmatrix}, \tag{2.101}$$

where n is the main principal quantum number, also known from non-relativistic quantum mechanics. We note that the inclusion of the main principal number n may seem arbitrary at this point, but it will become clearer in the next section.

Now it is time to simplify the Dirac Hamiltonian from (2.61)

$$H = -ic\vec{\alpha}\cdot\vec{\nabla} + U + c^2\beta, \tag{2.102}$$

where $U = -V$ is the electric potential energy. As $\vec{p} = -i\vec{\nabla}$, (2.102) can also be written as

$$\begin{aligned} H_D &= c\vec{\alpha}\cdot\vec{p} + U + c^2\beta \\ &= c\begin{pmatrix} 0 & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & 0 \end{pmatrix} + U + c^2\begin{pmatrix} \mathbb{1}_{2\times2} & 0 \\ 0 & -\mathbb{1}_{2\times2} \end{pmatrix} \end{aligned} \tag{2.103}$$

We have to derive how $\vec{\sigma}\cdot\vec{p}$ act on a relativistic state ψ. $\vec{\sigma}\cdot\vec{p}$ can be simplified to (see appendix A.1.9 for derivation).[32]

$$\begin{aligned} \vec{\sigma}\cdot\vec{p} &= \frac{1}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\Big(-ir\frac{\partial}{\partial r} + i\vec{\sigma}\cdot\vec{L}\Big) \\ &= \frac{1}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\Big(-ir\frac{\partial}{\partial r} + i\big(-K-1\big)\Big) \end{aligned} \tag{2.104}$$

$\frac{\vec{\sigma}\cdot\vec{x}}{r}$ needs to be investigated in some more detail. Firstly, we see that it changes sign under parity transformation and therefore it changes the parity of the state it acts upon. The parity of the spherical harmonics Y_{l,m_l} in (2.100) is given by $\big(-1\big)^l$[27]. If $\frac{\vec{\sigma}\cdot\vec{x}}{r}$ changes parity, it must also change the angular momentum quantum number l. Following the same trail of thoughts, it changes κ to $-\kappa$ and $-\kappa$ to κ,

$$\begin{aligned} \frac{\vec{\sigma}\cdot\vec{x}}{r}X_{\kappa,j,m_j} &= C_1 X_{-\kappa,j,m_j} \\ \frac{\vec{\sigma}\cdot\vec{x}}{r}X_{-\kappa,j,m_j} &= C_2 X_{\kappa,j,m_j}, \end{aligned} \tag{2.105}$$

where C_1 and C_2 are some constants. Although,

$$\left(\frac{\vec{\sigma}\cdot\vec{x}}{r}\right)^2 = 1, \tag{2.106}$$

which means that $C_1 = C_2 = \pm 1$. We choose to work with the convention $C_1 = C_2 = -1$.

$$\begin{aligned}\frac{\vec{\sigma}\cdot\vec{x}}{r}\chi_{\kappa,j,m_j} &= -X_{-\kappa,j,m_j}\\ \frac{\vec{\sigma}\cdot\vec{x}}{r}\chi_{-\kappa,j,m_j} &= -X_{\kappa,j,m_j},\end{aligned} \tag{2.107}$$

We can now solve the time-independent Dirac equation;

$$\begin{gathered}H(\vec{r})\psi(\vec{r}) = E\psi(\vec{r})\\ \Leftrightarrow\\ \left(c\begin{pmatrix}0 & \vec{\sigma}\cdot\vec{p}\\ \vec{\sigma}\cdot\vec{p} & 0\end{pmatrix} + U\mathbb{1}_{4\times4} + c^2\begin{pmatrix}\mathbb{1}_{2\times2} & 0\\ 0 & -\mathbb{1}_{2\times2}\end{pmatrix}\right)\frac{1}{r}\begin{pmatrix}P_{n,\kappa}(r)X_{\kappa,j,m_j}\\ iQ_{n,\kappa}(r)X_{-\kappa,j,m_j}\end{pmatrix} = E\frac{1}{r}\begin{pmatrix}P_{n,\kappa}(r)X_{\kappa,j,m_j}\\ iQ_{n,\kappa}(r)X_{-\kappa,j,m_j}\end{pmatrix},\end{gathered} \tag{2.108}$$

which can be converted into two coupled equations

$$\begin{aligned}c\vec{\sigma}\cdot\vec{p}\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j} + \left(U+c^2\right)\frac{1}{r}P_{n,\kappa}X_{\kappa,j,m_j} &= E\frac{1}{r}P_{n,\kappa}X_{\kappa,j,m_j}\\ c\vec{\sigma}\cdot\vec{p}\frac{1}{r}Q_{n,\kappa}X_{\kappa,j,m_j} + \left(U-c^2\right)\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j} &= E\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j}\end{aligned} \tag{2.109}$$

It is transformed by usage of (2.104)

$$\begin{aligned}\left(\frac{c}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\left(-ihr\frac{\partial}{\partial r} + i\left(-K-1\right)\right)\right)\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j} + \left(U+c^2\right)\frac{1}{r}P_{n,\kappa}X_{\kappa,j,m_j} &= E\frac{1}{r}P_{n,\kappa}X_{\kappa,j,m_j}\\ \left(\frac{c}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\left(-ihr\frac{\partial}{\partial r} + i\left(-K-1\right)\right)\right)\frac{1}{r}Q_{n,\kappa}X_{\kappa,j,m_j} + \left(U-c^2\right)\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j} &= E\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j}\end{aligned} \tag{2.110}$$

It is simplified by using (2.93) and multiplying with r from the left:

$$\begin{aligned}\left(c\frac{\vec{\sigma}\cdot\vec{x}}{r}\left(-ihr\frac{\partial}{\partial r} + i\left(\kappa-1\right)\right)\right)\frac{i}{r}Q_{n,\kappa}X_{-\kappa,j,m_j} + \left(U+c^2\right)P_{n,\kappa}X_{\kappa,j,m_j} &= EP_{n,\kappa}X_{\kappa,j,m_j}\\ \left(c\frac{\vec{\sigma}\cdot\vec{x}}{r}\left(-ihr\frac{\partial}{\partial r} + i\left(-\kappa-1\right)\right)\right)\frac{1}{r}Q_{n,\kappa}X_{\kappa,j,m_j} + \left(U-c^2\right)Q_{n,\kappa}X_{-\kappa,j,m_j} &= iEQ_{n,\kappa}X_{-\kappa,j,m_j}\end{aligned} \tag{2.111}$$

We can use the relations from (2.107) and the fact that $\frac{\vec{\sigma}\cdot\vec{x}}{r}$ commutes with $\frac{1}{r}$, as $\frac{\vec{\sigma}\cdot\vec{x}}{r}$ only acts on

$X_{\pm\kappa,j,m_j}$.

$$\begin{aligned} c\Big(-ir\frac{\partial}{\partial r}+i\big(\kappa-1\big)\Big)\frac{i}{r}Q_{n,\kappa}X_{\kappa,j,m_j}+\big(U+c^2\big)P_{n,\kappa}X_{\kappa,j,m_j}&=EP_{n,\kappa}X_{\kappa,j,m_j}\\ c\Big(-ir\frac{\partial}{\partial r}+i\big(-\kappa-1\big)\Big)\frac{1}{r}Q_{n,\kappa}X_{\kappa,j,m_j}+\big(U-c^2\big)Q_{n,\kappa}X_{-\kappa,j,m_j}&=iEQ_{n,\kappa}X_{-\kappa,j,m_j}\end{aligned} \tag{2.112}$$

We multiply the first line by X_{κ,j,m_J} and integrate over the angular variables $\phi\in[0,2\pi]$ and $\theta\in[0,\pi]$ in spherical coordinates. It should yield unity due to normalization. The second line is multiplied by $X_{-\kappa,j,m_j}$ and integrated over the angular variables as well. It should also yield unity due to normalization.

$$\begin{aligned} c\Big(-ir\frac{\partial}{\partial r}+i\big(\kappa-1\big)\Big)\frac{i}{r}Q_{n,\kappa}+\big(U+c^2\big)P_{n,\kappa}&=EP_{n,\kappa}\\ c\Big(-ir\frac{\partial}{\partial r}+i\big(-\kappa-1\big)\Big)\frac{1}{r}Q_{n,\kappa}+\big(U-c^2\big)Q_{n,\kappa}&=iEQ_{n,\kappa},\end{aligned} \tag{2.113}$$

which simplifies to

$$\begin{aligned} \Big(U+c^2\Big)P_{n,\kappa}+c\Big(\frac{\kappa}{r}-\frac{\partial}{\partial r}\Big)Q_{n,\kappa}&=EP_{n,\kappa}\\ \Big(U-c^2\Big)Q_{n,\kappa}+c\Big(\frac{\kappa}{r}+\frac{\partial}{\partial r}\Big)P_{n,\kappa}&=EQ_{n,\kappa}\end{aligned} \tag{2.114}$$

Equation (2.114) is known as the radial Dirac equation and needs to be solved.

2.7 Solution of the Radial Dirac equation

Let

$$\psi(\vec{r})=\frac{1}{r}\begin{bmatrix}P_\kappa(r)X_{\kappa,m_j}\\ iQ_\kappa(r)X_{-\kappa,m_j}\end{bmatrix} \tag{2.115}$$

The Dirac equation for hydrogenic atoms simplifies to, derived in the previous section,

$$\begin{aligned} \big[U(r)+c^2-E\big]P_k(r)+c\left[\frac{\kappa}{r}-\frac{d}{dr}\right]Q_k(r)&=0\\ \big[U(r)-c^2-E\big]Q_k(r)+c\left[\frac{\kappa}{r}+\frac{d}{dr}\right]P_k(r)&=0\end{aligned} \tag{2.116}$$

Where the boundary conditions are given as

$$\psi_{rad}(r)\equiv\begin{pmatrix}P_\kappa(r)\\ Q_\kappa(r)\end{pmatrix}=\begin{cases}0,\text{as } r\to 0\\ 0,\text{as } r\to\infty\end{cases}$$

and $U(r)$ is the Coulomb potential

$$U(r) = -\frac{Z}{r}$$

We analyze the dominating terms as r $\to \infty$ in equation (2.116):

$$\begin{aligned}(c^2 - E)P_\kappa(r) - c\frac{d}{dr}Q_\kappa(r) = 0 \\ (-c^2 - E)Q_\kappa(r) + c\frac{d}{dr}P_\kappa(r) = 0\end{aligned} \tag{2.117}$$

which transforms to the second order differential equation[33]

$$\frac{d^2}{dr^2}P_\kappa(r) = \left(c^2 - \frac{E^2}{c^2}\right)P_\kappa(r), \tag{2.118}$$

with the solution

$$P_\kappa(r) = \exp(\pm\lambda r),$$

where

$$\lambda = \sqrt{c^2 - \frac{E^2}{c^2}} \quad \in \quad \mathbb{R},$$

which restricts $|E| < c^2$. An electron's rest energy is c^2, which states that we search for a bound electron. Note that this approach will not solve the radial Dirac equation (2.116) for negative energy solutions, as that demands $E < -c^2$.

The only physically acceptable solution due to normalizability is

$$P_\kappa(r) = \exp(-\lambda r)$$

By letting

$$P_\kappa(r) = \sqrt{1 + \frac{E}{c^2}}\exp(-\lambda r),$$

it can be shown in equation (2.117) that

$$Q_\kappa(r) = -\sqrt{1 - \frac{E}{c^2}}\exp(-\lambda r)$$

To solve equation (2.116) for $r \to 0$, let

$$P_\kappa(r) = \sqrt{1+\frac{E}{c^2}}\exp(-\lambda r)(F_1 - F_2) \tag{2.119}$$

$$Q_\kappa(r) = \sqrt{1-\frac{E}{c^2}}\exp(-\lambda r)(F_1 + F_2) \tag{2.120}$$

We this back into equation (2.116),

$$\begin{aligned}\left[-\frac{Z}{r}+c^2-E\right]\sqrt{1+\frac{E}{c^2}}\exp(-\lambda r)(F_1-F_2)+c\left[\frac{\kappa}{r}-\frac{d}{dr}\right]\sqrt{1-\frac{E}{c^2}}\exp(-\lambda r)(F_1+F_2)=0\\ \left[-\frac{Z}{r}-c^2-E\right]\sqrt{1-\frac{E}{c^2}}\exp(-\lambda r)(F_1+F_2)+c\left[\frac{\kappa}{r}+\frac{d}{dr}\right]\sqrt{1+\frac{E}{c^2}}\exp(-\lambda r)(F_1-F_2)=0\end{aligned} \tag{2.121}$$

This simplifies to

$$\frac{d}{dx}F_1(x) = F_1(x)\left(-\frac{Z\nu}{xc}+\frac{c\nu}{2\lambda}-\frac{E\nu}{2\lambda c}+\frac{\kappa}{x}+\frac{1}{2}\right)+F_2(x)\left(\frac{Z\nu}{xc}-\frac{c\nu}{2\lambda}+\frac{E\nu}{2\lambda c}+\frac{\kappa}{x}+\frac{1}{2}\right)-\frac{d}{dx}F_2(x) \tag{2.122}$$

$$\frac{d}{dx}F_2(x) = F_1(x)\left(-\frac{Z}{xc\nu}-\frac{c}{2\lambda\nu}-\frac{E}{2\lambda c\nu}+\frac{\kappa}{x}-\frac{1}{2}\right)+F_2(x)\left(-\frac{Z}{xc\nu}-\frac{c}{2\lambda\nu}-\frac{E}{2\lambda c\nu}-\frac{\kappa}{x}+\frac{1}{2}\right)+\frac{d}{dx}F_1(x), \tag{2.123}$$

where

$$\nu = \frac{\sqrt{1+\frac{E}{c^2}}}{\sqrt{1-\frac{E}{c^2}}} \text{ and } x = 2\lambda r$$

We insert equation (2.123) into equation (2.122) and it yields

$$\begin{aligned}2\frac{d}{dx}F_1(x) =& F_1(x)\left(-\frac{Z}{xc}\left(\nu-\frac{1}{\nu}\right)+\frac{c}{2\lambda}\left(\nu+\frac{1}{\nu}\right)-\frac{E}{2\lambda c}\left(\nu-\frac{1}{\nu}\right)+1\right)+\\ & F_2(x)\left(\frac{Z}{xc}\left(\nu+\frac{1}{\nu}\right)-\frac{c}{2\lambda}\left(\nu-\frac{1}{\nu}\right)+\frac{E}{2\lambda c}\left(\nu+\frac{1}{\nu}\right)+\frac{2\kappa}{x}\right)\end{aligned} \tag{2.124}$$

It can be shown that $\left(\nu+\frac{1}{\nu}\right)=\frac{2c}{\lambda}$ and $\left(\nu-\frac{1}{\nu}\right)=\frac{2E}{c\lambda}$, which lead to

$$\frac{d}{dx}F_1(x) = F_1(x)\left(-\frac{ZE}{c^2\lambda x}+1\right)+F_2(x)\left(\frac{Z}{\lambda x}+\frac{\kappa}{x}\right) \tag{2.125}$$

Inserting equation (2.125) back into equation (2.123) yields

$$\frac{d}{dx}F_2(x) = F_1(x)\left(-\frac{Z}{\lambda x}+\frac{\kappa}{x}\right)+F_2(x)\left(\frac{ZE}{c^2\lambda x}\right) \tag{2.126}$$

We propose solutions of the form $F_1(x) = a_1 x^\gamma$ and $F_2(x) = a_2 x^\gamma$ as $r \to 0$.
When these solutions are inserted into equation (2.125) and equation (2.125), one can analyze the dominating terms as $r \to 0$,

$$\begin{aligned} \gamma a_1 &= \left(-\frac{ZE}{\lambda c^2}\right) a_1 + \left(\frac{Z}{\lambda} + \kappa\right) a_2 \\ \gamma a_2 &= \left(-\frac{Z}{\lambda} + \kappa\right) a_1 + \frac{ZE}{\lambda c^2 x} a_2 \end{aligned} \tag{2.127}$$

Factoring for $\frac{a_1}{a_2}$ yields

$$\frac{a_1}{a_2} = \frac{\kappa + \frac{Z}{\lambda}}{\gamma + \frac{ZE}{c^2\lambda}} = \frac{\gamma - \frac{ZE}{c^2\lambda}}{\kappa - \frac{Z}{\lambda}}, \tag{2.128}$$

which is equivalent to

$$\gamma^2 = \kappa^2 - \frac{Z^2}{c^2} \tag{2.129}$$

$\gamma = \sqrt{\kappa^2 - \frac{Z^2}{c^2}}$ is the only allowed physical solution as $r \to 0$, due to boundary conditions. The next task is to solve equation (2.125) and (2.126) for all $r \in \mathbb{R}$. Firstly, we express $F_1(x)$ in terms of $F_2(x)$ and $\frac{d}{dx}F_2(x)$ in equation (2.126),

$$F_1(x) = \frac{1}{\kappa - \frac{Z}{\lambda}} \left(x \frac{d}{dx} F_2(x) - \frac{ZE}{c^2\lambda} F_2(x) \right) \tag{2.130}$$

and insert this into equation (2.125) and yield a second order differential equation

$$x \frac{d^2}{dx^2} F_2(x) + (1-x) \frac{d}{dx} F_2(x) - \left(\frac{\gamma^2}{x} - \frac{ZE}{c^2\lambda} \right) F_2(x) = 0 \tag{2.131}$$

Letting

$$F_2(x) = x^\gamma F \tag{2.132}$$

will simplify equation (2.131) to Krummer's equation[33]

$$x \frac{d^2}{dx^2} F_2(x) + (b-x) \frac{d}{dx} F_2(x) - a F_2(x) = 0 \tag{2.133}$$

$$\text{where } b = 2\gamma + 1 \text{ and } a = \gamma - \frac{ZE}{c^2\lambda}$$

The solutions to Krummer's eq. (2.133) are the confluent hypergeoemetric functions,[33]

$$\begin{aligned} F(a,b,x) =& 1+\frac{a}{b}x+\frac{a(a+1)}{b(b+1)}\frac{x^2}{2!}+\frac{a(a+1)(a+2)}{b(b+1)(b+2)}\frac{x^3}{3!} \\ &+\cdots++\frac{a(a+1)(a+2)\cdots(a+k-1)}{b(b+1)(b+2)\cdots(b+k-1)}\frac{x^k}{k!}+\cdots \end{aligned} \tag{2.134}$$

Therefore, $F_2(x) = x^\gamma F(a,b,x)$, and we insert this relation into equation (2.130),

$$F_1(x) = \frac{x^\gamma}{\kappa-\frac{Z}{\lambda}}\left(aF(a,b,x)+x\frac{d}{dx}F(a,b,c)\right), \tag{2.135}$$

which simplifies to

$$F_1(x) = \frac{ax^\gamma}{\kappa-\frac{Z}{\lambda}}F(a+1,b,x) \tag{2.136}$$

When analyzing the hypergeometric function in (2.134), it is clearly divergent as r$\rightarrow\infty$ as its x-dependence is of the form;

$$exp(x) = \sum_{i=0}^{\infty}\frac{x^i}{i!} \tag{2.137}$$

To avoid divergence, the series must be truncated at $a=-n_r$, where $n_r\in\mathbb{N}$, which leads to

$$a = \gamma-\frac{ZE}{c^2\lambda} = -n_r \tag{2.138}$$

Solving this for the energy E leads to

$$E_{n_r,\kappa} = \frac{c^2}{\sqrt{1+\frac{Z^2\alpha^2}{(n_r+\gamma)^2}}}, \tag{2.139}$$

where the energy $E_{n_r,\kappa}$ clearly is dependent of the quantum number κ and a new temporary quantum number n_r. Maclaurin expansion of (2.139) about $Z^2\alpha^2$ yields

$$E_{n_r,\kappa} = c^2-\frac{Z^2}{2(n_r+|\kappa|)^2}+\frac{\alpha^2Z^4}{2(n_r+|\kappa|)^3}\left(\frac{3}{4(n_r+|\kappa|)}-\frac{1}{|\kappa|}\right)+\cdots \tag{2.140}$$

where the first term is the electron' rest energy and the second term is of similar form to the energy solution (2.141) for the non-relativistic and hydrogenic Schrödinger equation[27]

$$E_n = -\frac{Z^2}{2n^2}, \tag{2.141}$$

where n is the main principal quantum number. Clearly, it seems like we can assume $n_r = n-|\kappa|$. The remaining terms in (2.140) are relativistic corrections. We insert $n_r = n-|\kappa|$ back into (2.140)

and (2.138):

$$E_{n,\kappa} = \frac{c^2}{\sqrt{1+\frac{Z^2\alpha^2}{(n-|\kappa|+\gamma)^2}}} \qquad (2.142)$$
$$a = |\kappa| - n$$

The unnormalized radial solutions (2.119) and (2.120) can now be written as

$$P_{n,\kappa}(r) = \sqrt{1+\frac{E}{c^2}}\exp(-\lambda r)(2\lambda r)^\gamma\left(\frac{|\kappa|-n}{\kappa-\frac{Z}{\lambda}}F(|\kappa|+1-n,2\gamma+1,2\lambda r) - F(|\kappa|-n,2\gamma+1,2\lambda r)\right)$$
$$Q_{n,\kappa}(r) = \sqrt{1-\frac{E}{c^2}}\exp(-\lambda r)(2\lambda r)^\gamma\left(\frac{|\kappa|-n}{\kappa-\frac{Z}{\lambda}}F(|\kappa|+1-n,2\gamma+1,2\lambda r) + F(|\kappa|-n,2\gamma+1,2\lambda r)\right), \qquad (2.143)$$

and they are fixed by the quantum number κ and the main principal quantum number n.

2.8 Dirac Hamiltonian with External Pulse

We have already been introduced to the time-dependent Dirac equation;

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \Big(-ic\alpha_i\nabla_i + c^2\beta\Big)\psi(\vec{r},t) \qquad (2.144)$$

Our next task is to include both the Coulomb potential $V(\vec{r})$ and an external pulse $A(\vec{r},t)$. We do so with the method of minimal coupling from section 2.5. We introduce a four-vector $A_\mu = [V(\vec{r}), -A_i(\vec{r},t)]$. The method of minimal coupling yields

$$\begin{aligned}\frac{\partial}{\partial t} &\to \frac{\partial}{\partial t} - iV(\vec{r}) \\ \nabla_i &\to \nabla_i + iA_i(\vec{r},t)\end{aligned}, \qquad (2.145)$$

which transforms the time-dependent Dirac equation to

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \Big(-ic\alpha_i\nabla_i + c^2\beta - V(\vec{r}) + c\alpha_i A_i(\vec{r},t)\Big)\psi(\vec{r},t) \qquad (2.146)$$

We denote

$$\begin{aligned}H_0(\vec{r}) &= -ic\alpha_i\nabla_i + c^2\beta - V(\vec{r}) \\ H_I(\vec{r},t) &= c\alpha_i A_i(\vec{r},t),\end{aligned} \qquad (2.147)$$

such that

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \Big(H_0(\vec{r}) + H_I(\vec{r},t)\Big)\psi(\vec{r},t) \qquad (2.148)$$

The Dirac Hamiltonian with external field H_D is time-dependent and does not have stationary solutions. Because both H_0 and H_I are of a considerable size, approximations are necessary. It is also quite complex and it calls for a numerical solution.

Chapter 3

Numerical Methods

We shall in this chapter give an overview of the methods used to numerically solve the time-dependent Dirac equation with an external pulse. In contrast to analytic methods, methods with "pen and paper", numerical methods must constraint itself by working on discrete domains. Analytic functions can be defined on continuous grids, ranging from $-\infty$ to ∞. This is not possible in numerical methods, as it would require infinite numerical power. One must therefore define functions on discrete grids $I = [a, b]$ with a limited amount of points between a and b. Ideally would one want results and solutions to be independent of these constrains, but instead one faces a new challenge: Replace infinities with discrete values, without suffering from insufficient numerical power and still have sufficient accuracy.

3.1 The Eigenvalue Problem

As one often encounters eigenvalue problems in quantum mechanics, it is beneficial to be able to solve these problems. A general eigenvalue problem has the form

$$Ax = \lambda x, \tag{3.1}$$

where A is a matrix, x is the eigenvector and λ is the eigenvalue. When A is of a significant size, approximations and numerical solutions are necessary. We propose that x is a linear combination of basis functions f,

$$x = \sum_{i=1}^{\infty} c_i f_i \tag{3.2}$$

We must truncate the series at some integer N, as is it impossible to work with infinities numerically,

$$x \approx \sum_{i=1}^{N} c_i f_i \tag{3.3}$$

We insert it back into the eigenvalue problem (3.1)

$$\sum_{i=1}^{L} c_i A f_i = \lambda \sum_{i=1}^{L} c_i f_i, \tag{3.4}$$

and multiply with f_j for some arbitrary $1 \leq j \leq L$ and integrate over the volume V,

$$\sum_{i=1}^{L} c_i \int_V d^3 r\, f_j A f_i = \lambda \sum_{i=1}^{L} c_i \int_V d^3 r\, f_j f_i \tag{3.5}$$

We make the abbreviations

$$\begin{aligned} A_{ji} &= \int_V d^3 r\, f_j A f_i \\ S_{ji} &= \int_V d^3 r\, f_j f_i, \end{aligned} \tag{3.6}$$

which transforms (3.5) to

$$\sum_{i=1}^{L} c_i A_{ji} = \lambda \sum_{i=1}^{L} c_i S_{ji} \tag{3.7}$$

This system of equations can be written in a matrix representation, which is easier to solve numerically,

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1L} \\ A_{21} & A_{22} & \cdots & A_{2L} \\ \vdots & \ddots & & \vdots \\ A_{L1} & A_{L2} & \cdots & A_{LL} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_L \end{bmatrix} = \lambda \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1L} \\ S_{21} & S_{22} & \cdots & S_{2L} \\ \vdots & \ddots & & \vdots \\ S_{L1} & S_{L2} & \cdots & S_{LL} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_L \end{bmatrix}, \tag{3.8}$$

which in shorthand notation becomes

$$Ac = \lambda Sc \tag{3.9}$$

Such problems are numerically solvable for programming languages, like MATLAB.

3.2 The Eigenvalue Problem for the Radial Dirac Equation Without an External Pulse

In chapter 2 section 2.6 we derived the time-independent Dirac equation for hydrogenic atoms:

$$\begin{aligned}\Big(U + c^2\Big)P_{n,\kappa} + c\Big(\frac{\kappa}{r} - \frac{\partial}{\partial r}\Big)Q_{n,\kappa} &= E_{n,\kappa}P_{n,\kappa} \\ \Big(U - c^2\Big)Q_{n,\kappa} + c\Big(\frac{\kappa}{r} + \frac{\partial}{\partial r}\Big)P_{n,\kappa} &= E_{n,\kappa}Q_{n,\kappa},\end{aligned} \tag{3.10}$$

where $P_{n,\kappa}$ and $Q_{n,\kappa}$ are the radial part of the wavefunction ψ_{n,κ,j,m_k}

$$\psi(r,\phi,\theta)_{n,\kappa,j,m_k} = \frac{1}{r}\begin{pmatrix} P_{n,\kappa}(r)X_{\kappa,j,m_j}(\phi,\theta) \\ iQ_{n,\kappa}X_{-\kappa,j,m_j}(\phi,\theta)\end{pmatrix} \tag{3.11}$$

It is here of interest to find these radial solutions $P_{n,\kappa}$ and $Q_{n,\kappa}$ and the energy $E_{n,\kappa}$ in (3.10), as these will be used to solve the Dirac equation for hydrogenic atoms with an external pulse, equation (2.148) in chapter 2, section 2.8. We note that (3.10) indeed is a differential equation, so boundary conditions are necessary. This is discussed in section 3.7. We rewrite (3.10) into a matrix equation.

$$\begin{pmatrix} U + c^2 & c\Big(\frac{\kappa}{r} - \frac{\partial}{\partial t}\Big) \\ U - c^2 & c\Big(\frac{\kappa}{r} + \frac{\partial}{\partial t}\Big)\end{pmatrix}\begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa}\end{pmatrix} = E_{n,\kappa}\begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa}\end{pmatrix}, \tag{3.12}$$

which in shorthand notation is written as

$$H_{D_{rad}}\psi_{rad} = E\psi_{rad} \tag{3.13}$$

It is an eigenvalue problem and we can use the same procedure as in 3.1. The next task is to choose a complete basis $\{f_i\}$ our eigenveictors ψ_{rad} can be expanded in;

$$\psi_{rad} \approx \sum_{i=1}^{2N} c_i f_i \tag{3.14}$$

3.3 Choosing a Basis for Solving the Radial Dirac Equation

It is of great significance to choose an appropriate basis $\{f_i\}$ for solving eigenvalue problems. If one chooses an inappropriate basis, a great number of basis functions are needed to yield results with sufficient precision, whereas an appropriate basis needs fewer basis functions to yield results with the same precision. Therefore one can save both time and numerical power by choosing an appropriate basis.

In section 3.2 we were introduced to the expansion

$$\psi_{rad}(r) \approx \sum_{i=1}^{2N} c_i f_i(r), \tag{3.15}$$

where N is the truncation factor. We want N to be as small as possible without giving up too much precision. We must also restrict our basis functions to be defined at some finite grid $r \in [0, R_{max}]$. It could be an advantage to look at a basis with piecewise defined functions, as one lets some basis functions $\{f_i\}$ for $1 < i \leq L_1$ define ψ_{rad} at some domain $r \in [0, R_1]$ and $\{f_j\}$ for $L_1 < j \leq L_2$ define ψ_{rad} on $r \in [R_1, R_2]$ and so on. In this way, one avoids that too many basis functions interfere on the whole domain $r \in [0, R_{max}]$.

In quantum mechanics a basis set $\{v_1, \ldots v_n\}$ should ideally be complete;

$$\sum_n v_n^*(x)\, v_n(x') = \delta(x - x') \tag{3.16}$$

An obvious choice for such a complete basis is sine functions, although it is more convenient to use B-splines, as B-splines are piecewise defined functions. We note that $2N$ in (3.15) must go to infinity for the basis to be complete, but this is a numerical impossibility. So we must settle for a basis that is as complete as possible.

3.4 B-splines

B-splines are piecewise defined polynomials and have shown themselves to be well suited for numerical purposes in quantum mechanics.[34]

Let us define an interval $I = [0, R_{max}]$, divided into k subintervals $I_i = [\xi_i, \xi_{i+1}]$, where the $(k+1)$

points $\{\xi_i\}$, called breakpoints, are strictly increasing

$$0 = \xi_1 < \xi_2 < \cdots < \xi_{k+1} = R_{max} \tag{3.17}$$

B-splines are somewhat similar to polynomials and can be defined by its order n. We define a new sequence of $(2n+k-1)$ points, called knots $\{t_i\}$. It is related to breakpoints $\{\xi_i\}$ in the following way:

$$\begin{aligned} t_1 = t_2 \quad &= \cdots = t_n = \xi_1 = 0 \\ t_{i+n-1} = \xi_i \quad &\qquad \text{for } 2 \le i \le k \\ t_{k+n-1} = t_{k+n} \quad &= \cdots = t_{k+2n-1} = \xi_{k+1} = R_{max} \end{aligned} \tag{3.18}$$

The definition of B-splines depend on these knots $\{t_i\}$, as B-splines for order $n = 1$ take the form[34]

$$B_i^1(x) = 1 \qquad \text{for } t_i \le x < t_{i+1}, \tag{3.19}$$

It is often sufficient to choose equidistant knotpoints t_i. An example of a first-order B-spline with equidistant knotpoints is given in figure (3.1).

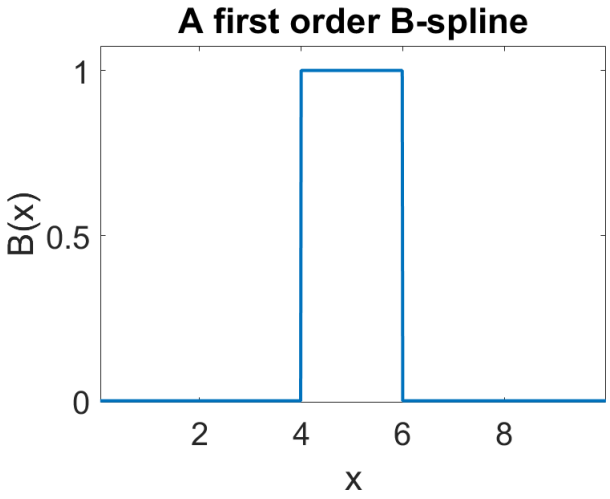


Figure 3.1: B-spline B_i with $t_i = 4$ and $t_{i+1} = 6$ defined on an interval $I = [0, 10]$.

A basis with first-order B-splines, defined on the same interval $I = [0, 10]$ as in figure (3.1), is shown in figure (3.2a). It becomes evident that no linear combination of the first order B-splines in figure (3.2a) can approximate the curvature of the wavefunction ψ well. It is therefore necessary to go to B-splines of higher order $n > 1$. Such higher order B-splines functions are defined through a recurrence relation[34]:

$$B_i^n(x) = \frac{x - t_i}{t_{i+n} - t_i} B_i^{n-1}(x) + \frac{t_{i+n+1} - x}{t_{i+n+1} - t_{i+1}} B_{i+1}^{n-1}(x), \tag{3.20}$$

where n is the order of the B-splines and $\{t_i\}$ are the same knotpoints one defined to yield the first order B-splines $\{B^1\}$. The higher order B-splines with $n > 1$ does show a disadvantage though. Different B-splines B_i^n are not necessarily orthogonal and may overlap, as one can see for $n = 2$ in figure (3.2b), $n = 3$ in figure (3.2c) and $n = 8$ in figure (3.2d).

It is noteworthy that the third order B-spline basis in figure (3.2c) is differentiable on the whole interval $I = [0, 10]$, unlike the first order B-spline basis in figure (3.2a) and the second order B-spline basis in figure (3.2b), which are only piecewise differentiable. It turns out that B-splines of order n is $(n-2)$ times differentiable on its entire domain.

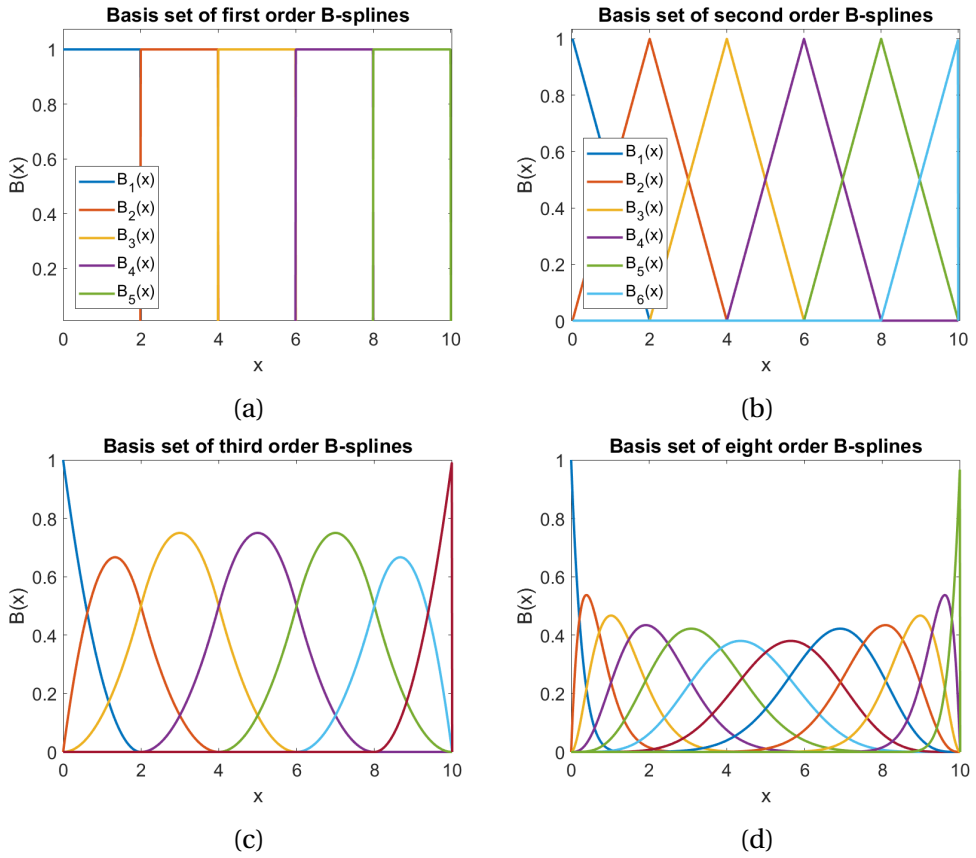


Figure 3.2: Various B-spline basis with different order n. $n = 1$ in (a), $n = 2$ in (b), $n = 3$ in (c) $n = 3$ and $n = 8$ in (d).

3.5 Application of B-splines

In section 3.2 it was stated that we want to expand the eigenvector $\psi_{rad} = \begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa} \end{pmatrix}$ in terms of basis functions $\{f_i\}$,

$$\psi_{rad} = \begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa} \end{pmatrix} \approx \sum_{i=1}^{2N} c_i f_i, \tag{3.21}$$

We have decided to use B-splines of order n as our basis set $\{B_i^n\}$. Therefore we propose for both the upper and lower component

$$\psi_{rad} = \begin{pmatrix} P_{n,\kappa} \\ Q_{n,-\kappa} \end{pmatrix} \approx \sum_{i=1}^{N} c_i B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{i=N+1}^{2N} c_i B_{i-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.22}$$

Thereafter we use the same procedure as in section 3.1, which yields

$$\begin{pmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,2N} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{2N,1} & H_{2N,2} & \cdots & H_{2N,2N} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2N} \end{pmatrix} = E \begin{pmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,2N} \\ S_{2,1} & S_{2,2} & \cdots & S_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{2N,1} & S_{2N,2} & \cdots & S_{2N,2N} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2N} \end{pmatrix}, \tag{3.23}$$

where

$$\begin{aligned} H_{i,j} &= \int_0^{R_{max}} dr \left(B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^\dagger H_{D_{rad}} B_j^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{for } i \le N, j \le N \\ H_{i,j} &= \int_0^{R_{max}} dr \left(B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^\dagger H_{D_{rad}} B_{j-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{for } i \le N, j > N \\ H_{i,j} &= \int_0^{R_{max}} dr \left(B_{i-1}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^\dagger H_{D_{rad}} B_j^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{for } i > N, j \le N \\ H_{i,j} &= \int_0^{R_{max}} dr \left(B_{i-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^\dagger H_{D_{rad}} B_{j-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{for } i > N, j > N, \end{aligned} \tag{3.24}$$

and

$$
\begin{aligned}
S_{i,j} &= \int_0^{R_{max}} dr \left(B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^\dagger B_j^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{for } i \leq N, j \leq N \\
S_{i,j} &= \int_0^{R_{max}} dr \left(B_{i-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^\dagger B_j^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 && \text{for } i > N, j \leq N \\
S_{i,j} &= \int_0^{R_{max}} dr \left(B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^\dagger B_{j-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 && \text{for } i \leq N, j > N \\
S_{i,j} &= \int_0^{R_{max}} dr \left(B_{i-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^\dagger B_{j-N}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{for } i > N, j > N
\end{aligned}
\tag{3.25}
$$

The integrals can be numerically calculated by using the Gauss-Legendre integration technique, which is a very precise integration method for polynomials. B-splines are indeed piecewise defined polynomials with minimal support, so it is an appropriate numerical integration technique. The eigenvalue problem in (3.23) can be calculated using the $eig()$-function in MATLAB.

We note that the matrix equation in (3.23) has $2N$ solutions, whereas the analytic solution would have infinite solutions. This loss of solutions is due to the truncation in (3.21). It is also noteworthy that all the $2N$ solutions will be bound states, due to the truncation of the r-grid at R_{max}.

The numerical solutions will instead be a distribution of some bound states and some pseudo continuum states. Pseudo continuum states are states which have energies in the energy continuum, both positive and negative as discussed in the next paragraph, but they are quantized both due to the boundary conditions and the fact that we work on a finite grid. It will be further discussed in section 3.7. We also note that we cannot fix states in either the positive or negative pseudo continuum with the main principle quantum number n, as they in fact are not bound. n from now on be denoted as an energy index pointing at the $2N$ solutions, but we will use a notation where $n = 1$ still is the ground state for simplicity.

Choosing a B-spline basis has a disadvantage, as it produces spurious states. These spurious states are not physical and stem from numerical limitations in handling high values. The eigenvalue solutions E, the energies, calculated from the eigenvalue problem (3.23) have a specific range. It can be in the negative energy continuum, $E \in (-\infty, -c^2)$, it can be in a bound state, $E \in (BE + c^2, c^2)$, where BE is the binding energy, and it can be in the positive energy continuum, $E \in (c^2, \infty)$. This is shown in figure (3.3).[35]

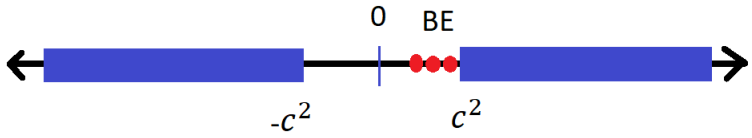


Figure 3.3: Range of the energies E from the radial Dirac equation. The thick blue line to the left is the negative energy continuum, the red dots are the bound states with binding energy BE, and the thick blue line to the right is the positive energy continuum.

As the energies can become quite large in absolute value, one encounters that the B-splines basis set is not sufficient to deal with this numerical challenge. The spurious states that arise are observed in the energy density in figure 3.4.

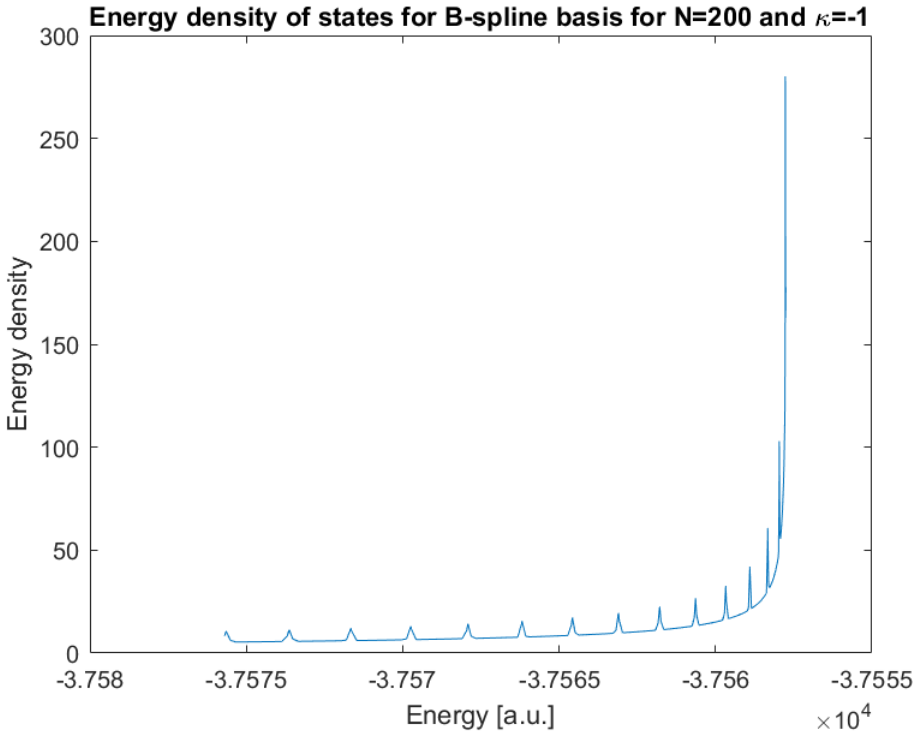


Figure 3.4: Energy density of the negative energy continuum for a B-spline basis, where $N = 200$, $\kappa = -1$ and $R_{max} = 100$.

All of the small tops in figure (3.4) indicate spurious states, as there is no physical reason for having multiple local maximums in the energy density anywhere in the negative energy continuum.

It has been suggested to expand $P_{n,\kappa}$ and $Q_{n,\kappa}$ in B-splines of different order[21], so we try

$$\psi_{rad} = \begin{pmatrix} P_{n,\kappa} \\ Q_{n,-\kappa} \end{pmatrix} = \sum_{i=1}^{N} c_i B_i^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{i=N+1}^{2N} c_i B_{i-N}^{n-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.26}$$

and we follow the same procedure as for equally ordered B-splines. It has shown itself to be an efficient method, as it does not yield the spurious states seen in figure 3.4. This is shown in figure 3.5.

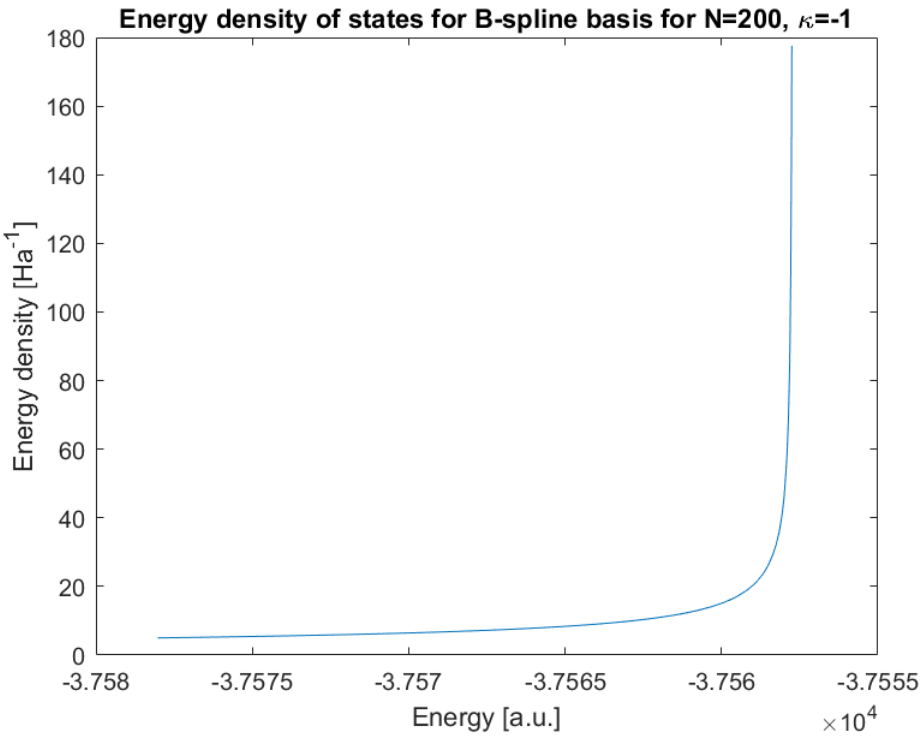


Figure 3.5: Energy density of the negative energy continuum for $E_{n,\kappa=-1}$ for a B-spline basis with different order n, where $N = 200$ and $R_{max} = 100$.

There are no local maximums in figure 3.5 and that indicates no spurious states. Although spurious states still occur for $\kappa > 0$.[36] It is shown for $\kappa = 1$, where state with the lowest energy of the bound states is spurious. Table (3.1) shows that the energy of the spurious state is much lower than theoretical expected energy. The spurious actually has lower energy then the theoretical ground state energy, which is a contradiction.

Table 3.1: Comparison of the energy of the spurious state, the energy of the first numerical excited state and theoretical expected energy. The computations are done with $\kappa = 1$, $N = 200$, $R_{max} = 100$. A B-spline basis with different orders has been used, where the upper component $P_{n,\kappa}$ is expanded in B-splines of eight order, and the lower component $Q_{n,\kappa}$ is expanded in B-splines of seventh order. The electrons rest mass has been subtracted from the energy.

	Theoretical expected energy	Energy of spurious state	Energy of first excited numerical state
$\kappa = 1$	-0.125002080189006	-15.522068605157983	-0.125002080196282

Table (3.1) strongly indicates that the first excited numerical state indeed is the physical bound state with the least energy for $\kappa = 1$. This is verified by figure 3.6, as the upper component of the theoretical bound state with least energy $P_{n=2,\kappa=1_{theory}}$ overlaps with the first excited numerical state $P_{n=3,\kappa=1_{num}}$.

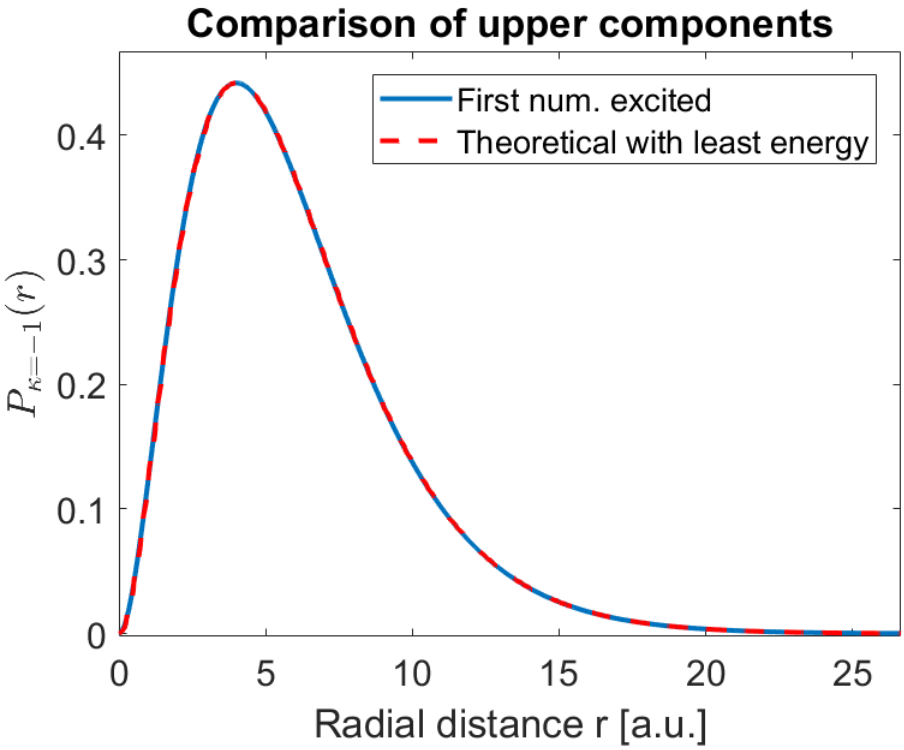


Figure 3.6: Comparison of the first excited numerical state $P_{3,1_{num}}$ with the theoretical bound state with least energy $P_{2,1_{theory}}$ for $\kappa = 1$, $N = 200$ and $R_{max} = 100$. A B-spline basis with different orders has been used, where the upper component P_κ is expanded in B-splines of eight order, and the lower component Q_κ is expanded in B-splines of seventh order.

The spurious state is on the other hand strongly oscillator near the nucleus and exhibits many nodes, which is seen in for the upper component in figure 3.7.

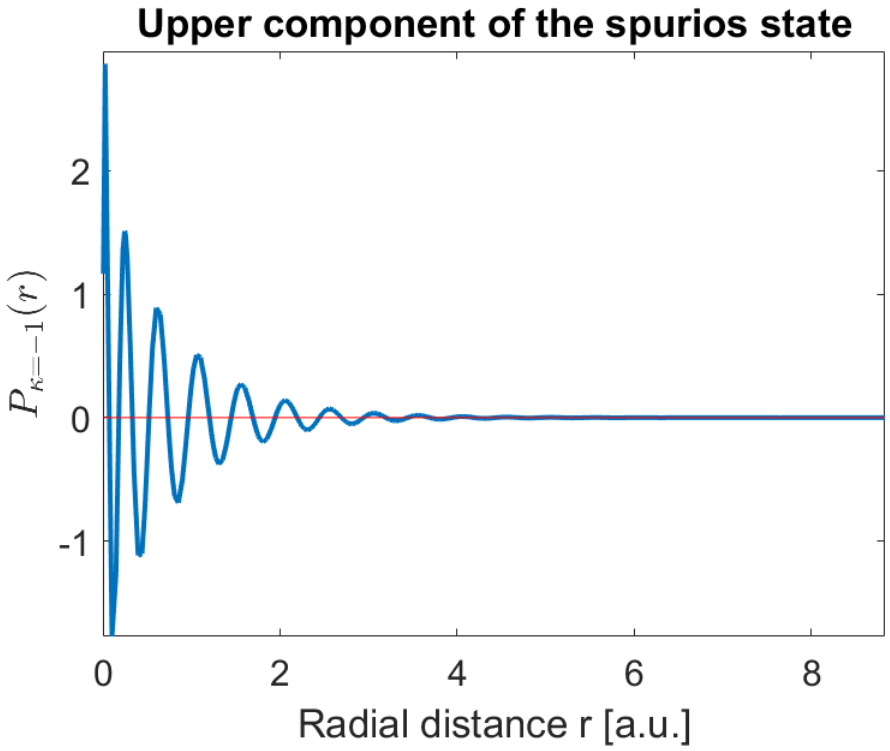


Figure 3.7: The upper component of the spurious state for $\kappa = 1$, $N = 200$, $R_{max} = 100$. A B-spline basis with different orders has been used, where the upper component $P_{n,\kappa}$ is expanded in B-splines of eight order, and the lower component $Q_{n,-\kappa}$ is expanded in B-splines of seventh order.

To proceed with numerical computations, we need to find another basis which does not produce spurious states. Such a basis is the dual kinetic balance, which is discussed in the next section.

3.6 Dual Kinetic Balance

The method of dual kinetic balance has the advantage that it does not yield the spurious states produced by the B-spline basis[36]. We let the eigenvector ψ_{rad} in the radial Dirac equation (3.12) in section 3.2 not be a simple sum of B-splines as in section 3.4, but instead we let

$$\psi_{rad} = \begin{pmatrix} P_{n,\kappa} \\ Q_{n,-\kappa} \end{pmatrix} = \sum_{i=1}^{N} c_i \begin{pmatrix} B_i^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_i^n + \frac{\kappa}{x}B_i^n\right) \end{pmatrix} + \sum_{i=N+1}^{2N} c_i \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{i-N}^n - \frac{\kappa}{x}B_{i-N}^n\right) \\ B_{i-N}^n \end{pmatrix}, \tag{3.27}$$

which yields the eigenvalue problem

$$\begin{pmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,2N} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{2N,1} & H_{2N,2} & \cdots & H_{2N,2N} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2N} \end{pmatrix} = E \begin{pmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,2N} \\ S_{2,1} & S_{2,2} & \cdots & S_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{2N,1} & S_{2N,2} & \cdots & S_{2N,2N} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2N} \end{pmatrix}, \tag{3.28}$$

where

$$\begin{aligned}
H_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} B_i^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_i^n + \frac{\kappa}{x}B_i^n\right) \end{pmatrix}^\dagger H_{D_{rad}} \begin{pmatrix} B_j^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_j^n + \frac{\kappa}{x}B_j^n\right) \end{pmatrix} && \text{for } i \le N, j \le N \\
H_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} B_i^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_i^n + \frac{\kappa}{x}B_i^n\right) \end{pmatrix}^\dagger H_{D_{rad}} \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{j-N}^n - \frac{\kappa}{x}B_{j-N}^n\right) \\ B_{j-N}^n \end{pmatrix} && \text{for } i \le N, j > N \\
H_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{i-N}^n - \frac{\kappa}{x}B_{i-N}^n\right) \\ B_{i-N}^n \end{pmatrix}^\dagger H_{D_{rad}} \begin{pmatrix} B_j^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_j^n + \frac{\kappa}{x}B_j^n\right) \end{pmatrix} && \text{for } i > N, j \le N \\
H_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{i-N}^n - \frac{\kappa}{x}B_{i-N}^n\right) \\ B_{i-N}^n \end{pmatrix}^\dagger H_{D_{rad}} \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{j-N}^n - \frac{\kappa}{x}B_{j-N}^n\right) \\ B_{j-N}^n \end{pmatrix} && \text{for } i > N, j > N,
\end{aligned} \tag{3.29}$$

and

$$\begin{aligned}
S_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} B_i^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_i^n + \frac{\kappa}{x}B_i^n\right) \end{pmatrix}^\dagger \begin{pmatrix} B_j^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_j^n + \frac{\kappa}{x}B_j^n\right) \end{pmatrix} && \text{for } i \le N, j \le N \\
S_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{i-N}^n - \frac{\kappa}{x}B_{i-N}^n\right) \\ B_{i-N}^n \end{pmatrix}^\dagger \begin{pmatrix} B_j^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_j^n + \frac{\kappa}{x}B_j^n\right) \end{pmatrix} && \text{for } i > N, j \le N \\
S_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} B_i^n \\ \frac{1}{2c}\left(\frac{d}{dr}B_i^n + \frac{\kappa}{x}B_i^n\right) \end{pmatrix}^\dagger \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{j-N}^n - \frac{\kappa}{x}B_{j-N}^n\right) \\ B_{j-N}^n \end{pmatrix} && \text{for } i \le N, j > N \\
S_{i,j} &= \int_0^{R_{max}} dr \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{i-N}^n - \frac{\kappa}{x}B_{i-N}^n\right) \\ B_{i-N}^n \end{pmatrix}^\dagger \begin{pmatrix} \frac{1}{2c}\left(\frac{d}{dr}B_{j-N}^n - \frac{\kappa}{x}B_{j-N}^n\right) \\ B_{j-N}^n \end{pmatrix} && \text{for } i > N, j > N
\end{aligned} \tag{3.30}$$

The integrals are again solved using the Gauss-Legendre integration technique and the eigenvalue problem is solved using $eig()$ in MATLAB. The method yields the energies E_n and the radial wavefunctions ψ_{rad}.

To verify the absence of spurious states, we take a closer look at the energy density in the negative continuum in figure 3.8.

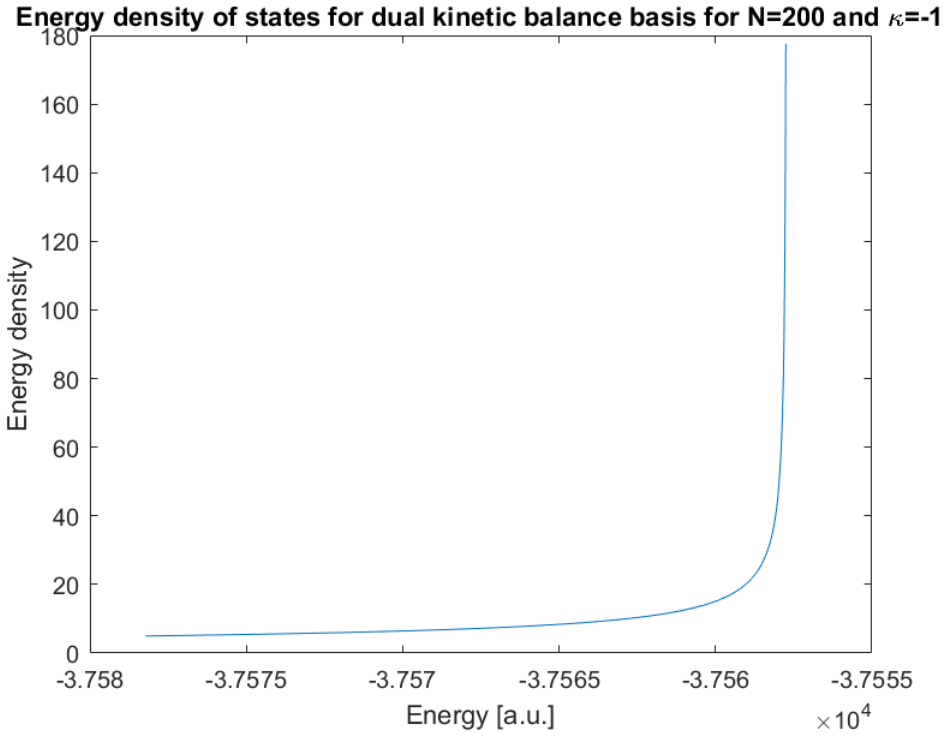


Figure 3.8: Energy density of the negative energy continuum for $E_{n,\kappa=-1}$ for a dual kinetic balance basis, where $N = 200$ and $R_{max} = 100$.

Figure 3.8 indicates no spurious states, as there are no local maximums. Spurious state does not appear for $\kappa > 0$ either, which it did for the method with B-splines of different order in table 3.1. We look at the numerical bound state with least energy for $\kappa = 1$, which shows great correspondence with theory. See table 3.2.

Table 3.2: Comparison of theoretical and numerical lowest bound state energies for $\kappa = 1$, $N = 200$ and $R_{max} = 200$. We note that the electron rest mass has been subtracted.

	Theoretical bound state with least energy	**Numeric bound state with least energy**
Energy	-0.125002080189006	-0.125002080189006

We have now found an appropriate basis to expand ψ_{rad} in, and it is time to determine the boundary conditions for ψ_{rad}.

3.7 Application and Choice of Boundary Conditions

3.7.1 Choosing Boundary Conditions

In section 3.2 we stated that the radial Dirac equation (3.31) is a differential equation and boundary conditions need to be determined.

$$\begin{pmatrix} U+c^2 & c\left(\frac{\kappa}{r}-\frac{\partial}{\partial t}\right) \\ U-c^2 & c\left(\frac{\kappa}{r}+\frac{\partial}{\partial t}\right) \end{pmatrix} \begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa} \end{pmatrix} = E_{n,\kappa} \begin{pmatrix} P_{n,\kappa} \\ Q_{n,\kappa} \end{pmatrix}, \tag{3.31}$$

which in shorthand notation is written as

$$H_{D_{rad}} \psi_{rad} = E \psi_{rad}, \tag{3.32}$$

and we have chosen $\psi_{rad}(r)$ to be defined for $r \in [0, R_{max}]$. Naturally, we must demand

$$\psi_{rad}(0) = 0, \tag{3.33}$$

as the electron must have zero probability to exist in the nucleus. It has been suggested to set $\psi_{rad}(R_{max}) = 0$[37, 36], which for bound states ψ_{rad} with energies $E \in (c^2 + BE, c^2)$ is a decent approximation. It is a decent approximation as the bound state ψ_{rad} must vanish for $R_{max} >> 1$, due to normalizability. However, we emphasize that the choice of R_{max} has great influence on how many and how well excited states are represented in the solutions.

The suggestion $\psi_{rad}(R_{max}) = 0$ is also a decent approximation for the states in the positive continuum with energies $E \in (c^2, \infty)$. It does, however, quantize the positive continuum. The physical continuum is naturally enough continuous and not quantized, which means that the suggestion creates a modified continuum, a pseudo continuum.

The boundary condition $\psi_{rad}(R_{max}) = 0$ will also create a pseudo continuum for states in the negative energy continuum with energies $E \in (-\infty, -c^2)$, although the case for negative continuum state is not as trivial as for positive continuum states. According to the Dirac sea theory, discussed in section 2.4, the states ψ_{rad} with energies in negative energy continuum are actually electrons in the Dirac sea. It has been shown that the boundary condition $\psi_{rad}(R_{max}) = 0$ cannot be applied in the Dirac equation for a free electron in a box[38]. It gives reason to suspect that by setting $\psi_{rad}(R_{max}) = 0$, one evokes unphysical behaviour for ψ_{rad} at $r = R_{max}$. This is seen in figure 3.9, where one has applied the boundary condition $\psi_{rad}(R_{max}) = 0$.

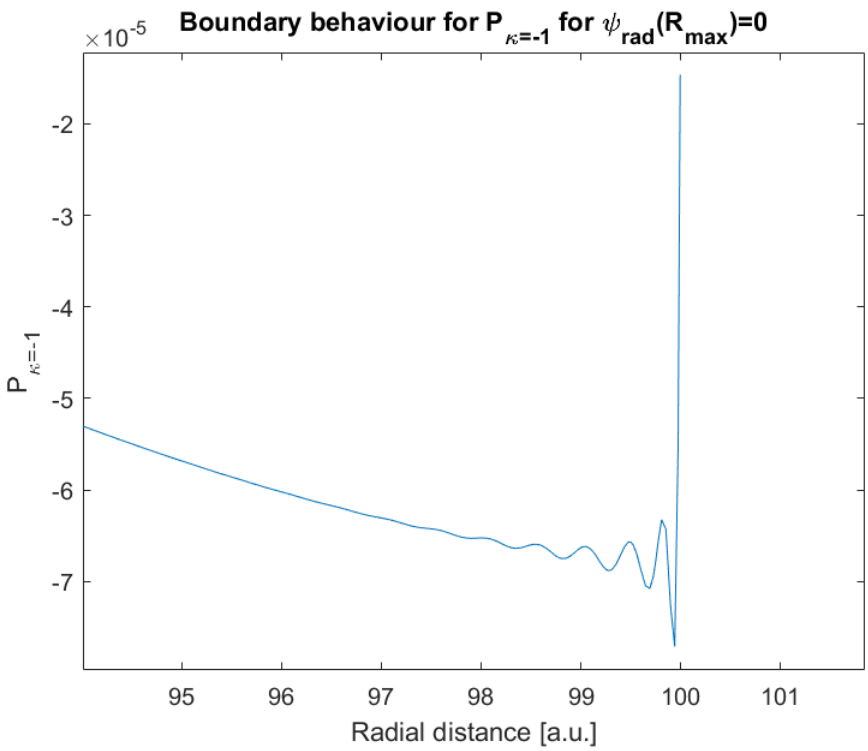


Figure 3.9: Boundary behaviour for $P_{\kappa=-1}$ when the boundary condition $\psi_{rad}(R_{max}) = 0$ is applied. It is numerically calculated with $\kappa = -1$, $N = 200$ and $R_{max} = 100$. The state with the highest energy in the negative energy continuum is plotted.

The oscillations in figure 3.9 seem unphysical, as it is no physical reason for this behaviour at the boundary. The other boundary choice at $r = R_{max}$, which seems more physical, is simply $\psi_{rad}(R_{max}) \neq 0$. The two different boundary conditions yield negligible difference for positive energy states, but some difference is seen for negative energy states. Firstly, the two different boundary conditions yields very similar $P_{\kappa=-1}$ for most of $r \in (0, R_{max})$. This is seen in figure 3.10.

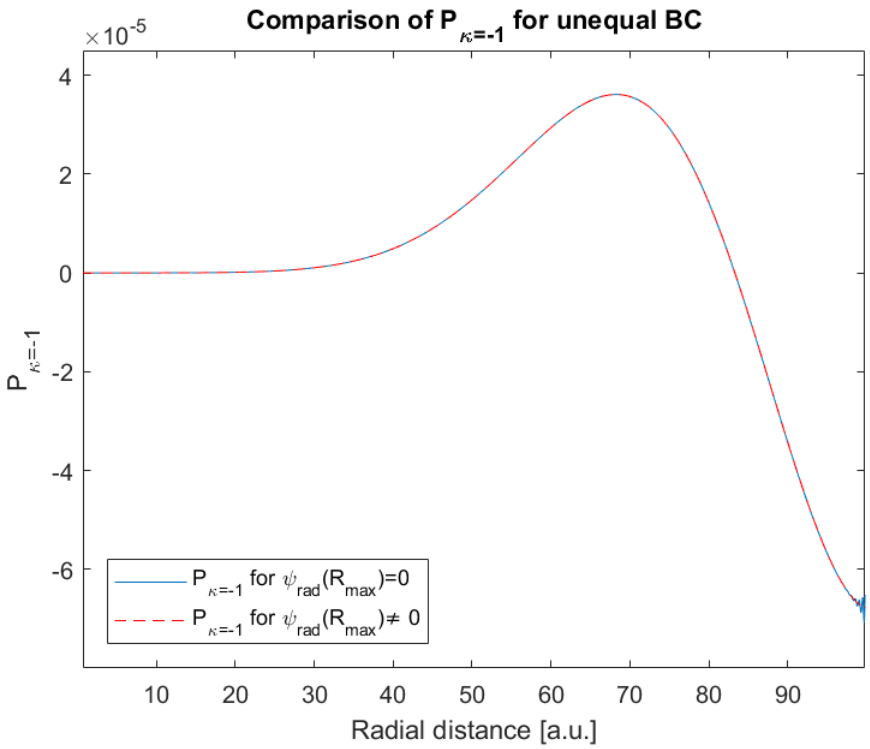


Figure 3.10: Comparison of $P_{\kappa=-1}$ for the two different boundary conditions $\psi_{rad}(R_{max}) \neq 0$ and $\psi_{rad}(R_{max}) = 0$. It is a comparison for the state with the lowest energy in the negative energy continuum with $\kappa = -1$ $N = 200$ and $R_{max} = 100$.

However, the two different boundary condition choices yield some difference at the boundary $r = R_{max}$, which is observed in figure 3.11.

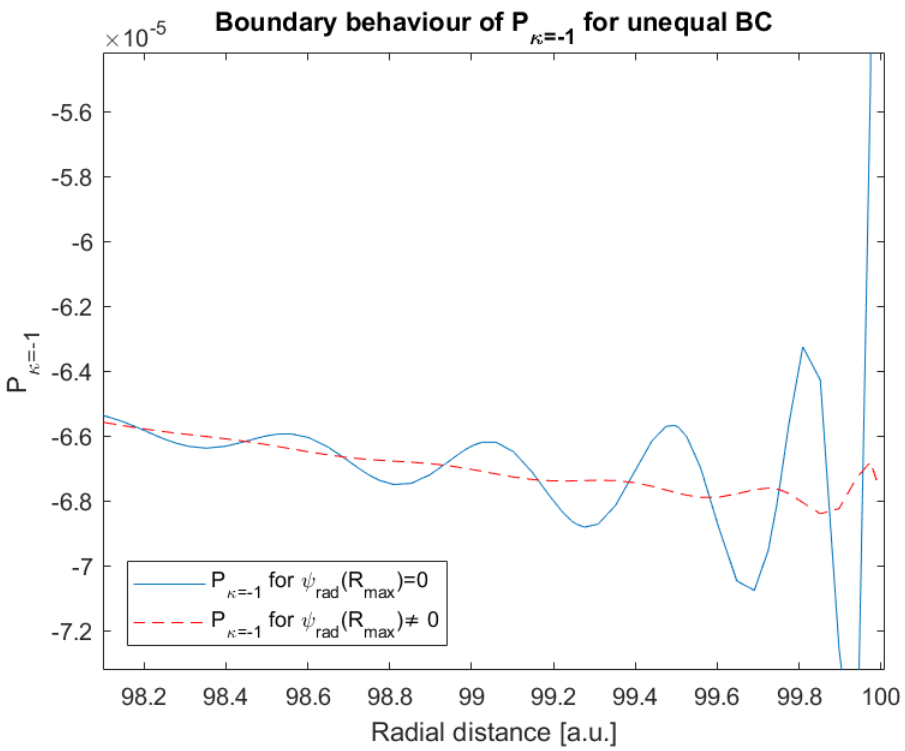


Figure 3.11: Comparison of $P_{\kappa=-1}$ for the two different boundary conditions $\psi_{rad}(R_{max}) \neq 0$ and $\psi_{rad}(R_{max}) = 0$ at the boundary $r \to R_{max}$. It is a comparison for the state with the lowest energy in the negative energy continuum for $\kappa = -1$ $N = 200$ and $R_{max} = 100$.

We have observed that both boundary condition choices yield some oscillation as $r \to R_{max}$, but $\psi_{rad}(R_{max}) \neq 0$ yields less, as seen in figure 3.11. Both boundary condition choices compute the energies for the lowest bound with good precision, but $\psi_{rad}(R_{max}) \neq 0$ seems like the better choice. So we proceed with

$$\psi_{rad}(R_{max}) \neq 0 \tag{3.34}$$

3.7.2 Application of Boundary Conditions

We have now established the following boundary conditions

$$\begin{aligned} \psi_{rad}(0) &= 0 \\ \psi_{rad}(R_{max}) &\neq 0 \end{aligned} \tag{3.35}$$

We apply $\psi_{rad}(0) = 0$ by letting all the basis functions $\{f_i\}$ in the dual kinetic balance basis set satisfy

$$f_i(0) = 0 \tag{3.36}$$

This is numerically incorporated by not expanding the upper component $P_{n,\kappa}$ and the lower component $Q_{n,\kappa}$ in their first respective dual kinetic balance basis functions f_{start}. f_{start} are the only basis functions in the basis set which are unequal zero at $r = 0$ and by removing them, one removes the possibility of $\psi_{rad}(0) \neq 0$.

$\psi_{rad}(R_{max}) \neq 0$ is applied by including the last respective basis functions f_{end} for the upper component $P_{n,\kappa}$ and the lower component $Q_{n,\kappa}$, as f_{end} are unequal zero at $r = R_{max}$,

$$f_{end}(R_{max}) \neq 0 \tag{3.37}$$

At the same time one wants the basis set to be as complete as possible. Removing basis functions from the basis set disrupts completeness. The most precise results have been calculated while removing the inner most B-spline at $r = 0$ and the outer most B-spline at $r = R_{max}$. An illustration of the chosen B-splines at the boundaries used in the dual kinetic balance basis is shown in figure 3.12 and figure 3.13.

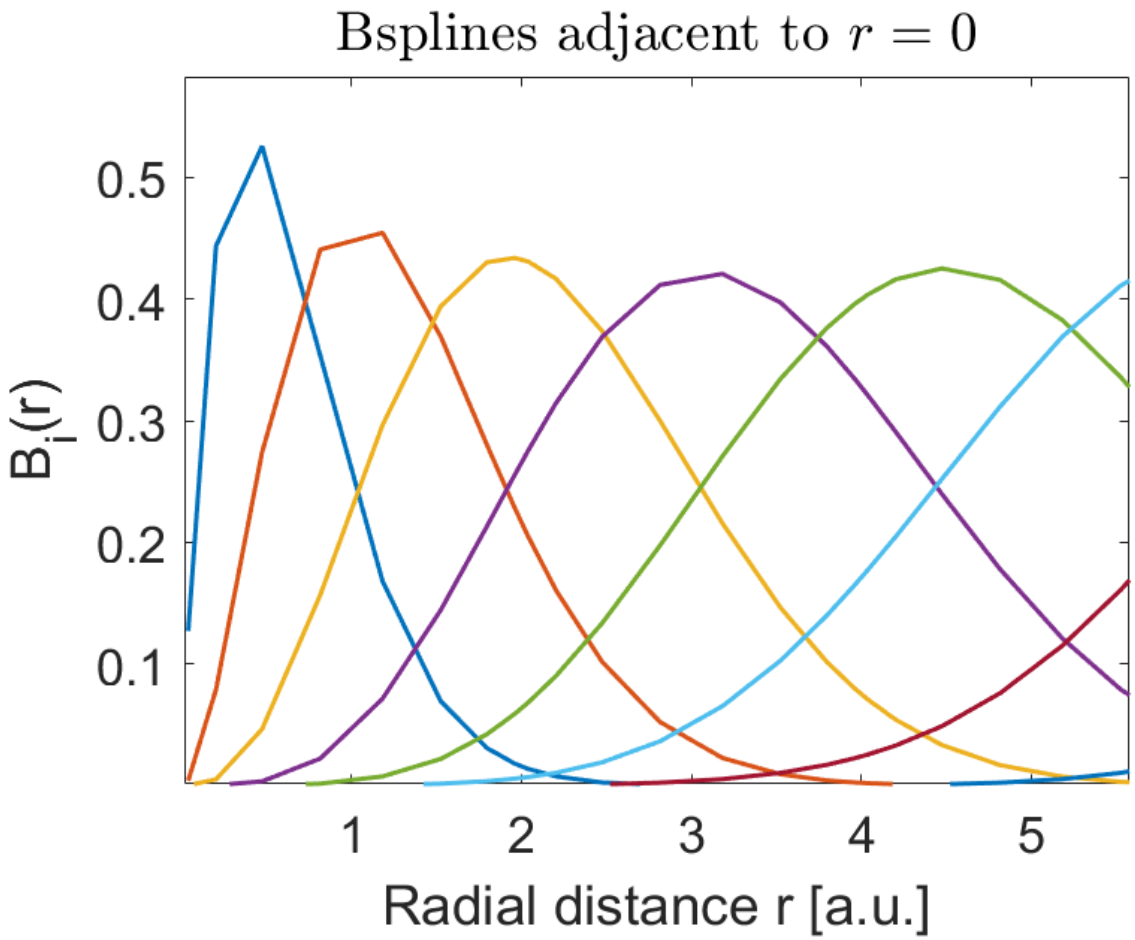


Figure 3.12: An illustration of the B-splines of eight order used in the dual kinetic balance basis adjacent to $r = 0$.

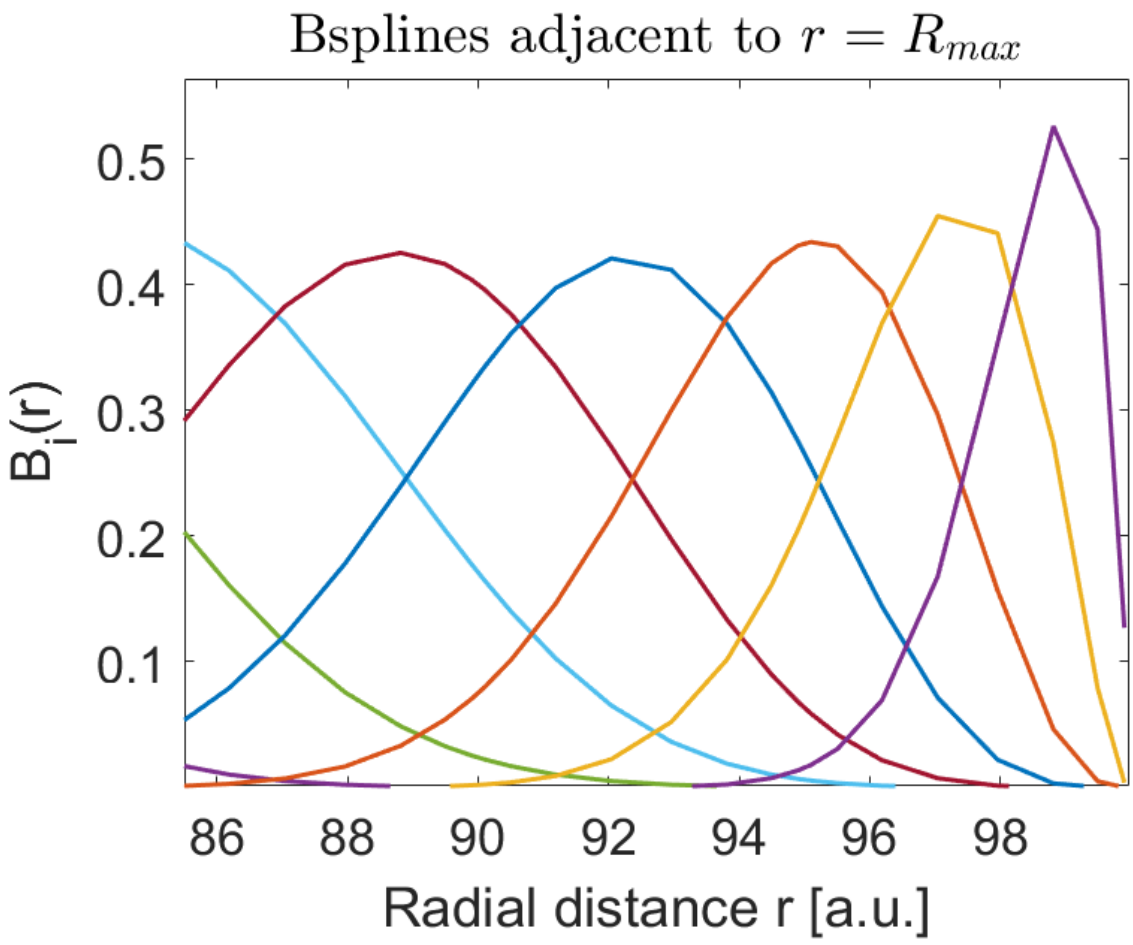


Figure 3.13: An illustration of the B-splines of eight order used in the dual kinetic balance basis adjacent to $r = R_{max}$.

We notice that the outermost and the inmost B-spline is near but unequal to zero near $r = 0$ and near $r = R_{max}$, which is a compromise between the boundary conditions in (3.35) and the fact that we want our basis set be as complete as possible.

3.8 Dirac Equation with External Pulse

The Dirac equation without an external pulse yielded stationary solutions, which lead to the already solved time-independent Dirac equation. When one wants to observe laser-atom interactions, one must naturally include an external pulse in the Dirac equation. The external pulse oscillates in time, and stationary solutions are no longer valid. New methods and techniques are necessary to solve the time-dependent Dirac equation (3.38, from chapter 2, section 2.8).

$$i\frac{\partial}{\partial t}\psi(\vec{r},t) = \Big(H_0(\vec{r}) + H_I(\vec{r},t)\Big)\psi(\vec{r},t), \tag{3.38}$$

where

$$\begin{aligned} H_0(\vec{r}) &= -ic\alpha_i\nabla_i + c^2\beta - V(\vec{r}) \\ H_I(\vec{r},t) &= c\alpha_i A_i(\vec{r},t), \end{aligned} \tag{3.39}$$

We have already solved the Dirac Hamiltonian without an external pulse, which yielded the stationary solutions

$$\psi_{n,\kappa,j,m_j} = \frac{1}{r}\begin{pmatrix} P_{n,\kappa}X_{\kappa,j,m_j} \\ iQ_{n,\kappa}X_{-\kappa,j,m_j} \end{pmatrix} \tag{3.40}$$

When solving (3.38), we want to take benefit from the stationary solutions. We propose solutions of the form[27]

$$\psi(\vec{r},t) = \sum_k c_k(t)\psi_k(\vec{r}), \tag{3.41}$$

where $\psi_k(\vec{r})$ are the stationary solutions from (3.40). The k-variable must sum over all m_j, κ and energy index n values. The summation over all κ values indicates summation over all j values, as $\kappa^2 = j+\frac{1}{2}$, which was already discussed in chapter 2, section 2.6. We state that the electron in our hydrogen atom will be in the ground state solution at $t=0$, where $\kappa=-1$ and $m_j=\frac{1}{2}$. It will later on be shown that the solutions of (3.38) are decoupled for different values $m_j = j,...,-j$, so there is no need to sum over all m_j-values. We also need to truncate the number kappa values we want to sum over, as we cannot work with infinities numerically. The number of n-values is truncated naturally as $n=2N$, where $2N$ is the number of basis functions.

We must also demand that the time-dependent wavefunction $\psi(\vec{r},t)$ is normalized at every time t,

$$1 = \int \psi^*(\vec{r},t)\psi(\vec{r},t)d^3r = \sum_m\sum_k c_m^* c_k \int \psi_m^*\psi_k d^3r = \sum_k c_k^* c_k, \tag{3.42}$$

which yields

$$\sum_k |c_k(t)|^2 = 1 \tag{3.43}$$

We insert the expansion of $\psi(\vec{r},t)$ from (3.41) into the time-dependent Dirac equation (3.38)

$$\sum_k i\frac{d}{dt}c_k(t)\psi_k(\vec{r}) = \sum_k c_k(t)\big(H_0(\vec{r}) + H_I(\vec{r},t)\big)\psi_k(\vec{r}) \tag{3.44}$$

How H_0 acts on $\psi_k(\vec{r})$ is already solved,

$$H_0(\vec{r})\psi_k(\vec{r}) = E_k\psi_k(\vec{r}) \tag{3.45}$$

By multiplying with $\psi_n^*(\vec{r})$ and integrate over the whole space, one yields

$$i\frac{d}{dt}c_n(t) = \sum_k c_k\Big(E_k\delta_{n,k} + \int \psi_n(\vec{r})H_I(\vec{r},t)\psi_k(\vec{r})d^3r\Big) \tag{3.46}$$

We define

$$V_{nk}(t) = \int \psi_n^*(\vec{r}) H_I(\vec{r},t)\psi_k(\vec{r}) d^3r, \tag{3.47}$$

which will be solved in section 3.9. This notation simplifies (3.46) to

$$i\frac{d}{dt}c_n(t) = \sum_k c_k(E_k\delta_{n,k} + V_{nk}) \tag{3.48}$$

This can be written in matrix form

$$i\frac{d}{dt}\begin{bmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_{k_{max}}(t) \end{bmatrix} = \begin{bmatrix} E_1+V_{11} & V_{12} & \dots & V_{1k_{max}} \\ V_{21} & E_2+V_{22} & \dots & V_{2k_{max}} \\ \vdots & \vdots & \vdots & \ddots \\ V_{k_{max}1} & V_{k_{max}2} & \dots & E_{k_{max}}+V_{k_{max}k_{max}} \end{bmatrix} \begin{bmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_{k_{max}}(t) \end{bmatrix} \tag{3.49}$$

The matrix system will in shorthand notation take the form of

$$\frac{d}{dt}\vec{c}(t) = H(t)\vec{c} \tag{3.50}$$

The system is firstly developed through the lowest order of the Magnus expansion by neglecting time-ordering[24]. Equation (3.49) simplifies to

$$\vec{c}(t+\Delta t) = \exp\Big(-i\int_t^{t+\Delta t} H(t)dt\Big)\vec{c(t)} \tag{3.51}$$

We use Riemann integration as $\Delta t << 1$

$$\vec{c}(t+\Delta t) = \exp\Big(-iH(t)\Delta t\Big)\vec{c(t)} \tag{3.52}$$

We simplify the expression by using the Maclaurin series for an exponential, as it is not trivial how to deal with operators in exponentials,

$$\begin{aligned} \vec{c}(t+\Delta t) &= \Big(\sum_{n=0}^{\infty}\frac{\big(-iH(t)\Delta t\big)^n}{n!}\Big)\vec{c(t)} \\ &= \sum_{n=0}^{\infty}\frac{\big(-i\Delta t\big)^n}{n!}H^n(t)\vec{c(t)} \end{aligned} \tag{3.53}$$

The numerical calculation of $H(t)^n$ takes a lot of numerical power and it needs to be approximated. It is solved by using Krylov subspace methods, which is discussed in section 3.10.

3.9 Numerical Approach for Solving the Interaction Hamiltonian

In section 3.8 we were asked to solve

$$V_{mk}(t) = \int dr^3 \psi_m(\vec{r}) H_I(\vec{r},t) \psi_k(\vec{r}) = \langle \psi_{n,\kappa,j,m_j} | H_I | \psi_{n',\kappa',j',m'_j} \rangle, \tag{3.54}$$

where

$$H_I(\vec{r},t) = c\vec{\alpha} \cdot \vec{A}(\vec{r},t), \tag{3.55}$$

and $\vec{A}(\vec{r},t)$ is the vector potential of the external field. In the dipole approximation, one neglects the spatial part of the vector field $\vec{A}$;

$$\vec{A}(\vec{r},t) \approx \vec{A}(t) \tag{3.56}$$

If one assumes that the external vector field $\vec{A}$ is polarized along the z-direction, the vector field $\vec{A}$ simplifies to

$$\vec{A}(t) = A(t)\hat{z}, \tag{3.57}$$

which inserted back into (3.55) yields

$$\begin{aligned} H_I &= c\alpha_z A(t) \\ &= cA(t) \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \end{aligned} \tag{3.58}$$

We have previously derived that the wavefunction ψ_{n,κ,j,m_j} has the form

$$\psi_{n,\kappa,j,m_j} = \frac{1}{r} \begin{pmatrix} P_{n,\kappa}(r) X_{\kappa,j,m_j} \\ iQ_{n,\kappa} X_{-\kappa,j,m_j} \end{pmatrix}, \tag{3.59}$$

where

$$\begin{aligned} X_{\kappa,j,m_j} &= \sum_{m_s,m_l} \langle l_\kappa, m_{l_\kappa}, s = \frac{1}{2}, m_s | j, m_j \rangle Y_{l_\kappa,m_{l_\kappa}} \chi_{m_s} \\ X_{-\kappa,j,m_j} &= \sum_{m_s,m_l} \langle l_{-\kappa}, m_{l_{-\kappa}}, s = \frac{1}{2}, m_s | j, m_j \rangle Y_{l_{-\kappa},m_{l_{-\kappa}}} \chi_{m_s}, \end{aligned} \tag{3.60}$$

which simplifies to

$$\begin{aligned} X_{\kappa,j,m_j} &= \langle l_\kappa, m_j - \frac{1}{2}, s = \frac{1}{2}, \frac{1}{2} | j, m_j \rangle Y_{l_\kappa, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \langle l_\kappa, m_j + \frac{1}{2}, s = \frac{1}{2}, -\frac{1}{2} | j, m_j \rangle Y_{l_\kappa, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ X_{-\kappa,j,m_j} &= \langle l_{-\kappa}, m_j - \frac{1}{2}, s = \frac{1}{2}, \frac{1}{2} | j, m_j \rangle Y_{l_{-\kappa}, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \langle l_{-\kappa}, m_j + \frac{1}{2}, s = \frac{1}{2}, -\frac{1}{2} | j, m_j \rangle Y_{l_{-\kappa}, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}, \tag{3.61}$$

It can in shorthand notation be written as

$$\begin{aligned} X_{\kappa,j,m_j} &= C^{j,l_\kappa}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C^{j,l_\kappa}_{m_j + \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ X_{-\kappa,j,m_j} &= C^{j,l_{-\kappa}}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_{-\kappa}, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C^{j,l_{-\kappa}}_{m_j + \frac{1}{2}, -\frac{1}{2}} Y_{l_{-\kappa}, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}, \tag{3.62}$$

Let us now inspect how H_I acts on the wavefunction ψ_{n,κ,j,m_j}

$$\begin{aligned} H_I \psi_{n,\kappa,j,m_j} &= cA(t) \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \frac{1}{r} \begin{pmatrix} P_{n,\kappa}(r) X_{\kappa,j,m_j} \\ iQ_{n,\kappa}(r) X_{-\kappa,j,m_j} \end{pmatrix} \\ &= \frac{cA(t)}{r} \begin{pmatrix} iQ_{n,\kappa}(r) \sigma_z X_{-\kappa,j,m_j} \\ P_{n,\kappa}(r) \sigma_z X_{\kappa,j,m_j} \end{pmatrix} \\ &\overset{(3.62)}{=} \frac{cA(t)}{r} \begin{pmatrix} iQ_{n,\kappa}(r) \left(C^{j,l_{-\kappa}}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_{-\kappa}, m_j - \frac{1}{2}} \sigma_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C^{j,l_{-\kappa}}_{m_j + \frac{1}{2}, -\frac{1}{2}} Y_{l_{-\kappa}, m_j + \frac{1}{2}} \sigma_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ P_{n,\kappa}(r) \left(C^{j,l_\kappa}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j - \frac{1}{2}} \sigma_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C^{j,l_\kappa}_{m_j + \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j + \frac{1}{2}} \sigma_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \end{pmatrix} \\ &= \frac{cA(t)}{r} \begin{pmatrix} iQ_{n,\kappa}(r) \left(C^{j,l_{-\kappa}}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_{-\kappa}, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - C^{j,l_{-\kappa}}_{m_j + \frac{1}{2}, -\frac{1}{2}} Y_{l_{-\kappa}, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ P_{n,\kappa}(r) \left(C^{j,l_\kappa}_{m_j - \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j - \frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - C^{j,l_\kappa}_{m_j + \frac{1}{2}, \frac{1}{2}} Y_{l_\kappa, m_j + \frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \end{pmatrix} \end{aligned} \tag{3.63}$$

We multiply with a state $\psi^\dagger_{n',\kappa',j,m_j}$ from the left and integrate over the whole space $\mathscr{R}_3$. We note that we work in spherical coordinates, where

$$dV = r^2 sin\theta dr d\theta d\phi, \tag{3.64}$$

and $r \in [0,\infty)$, $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$. This transforms (3.63) to

$$
\begin{aligned}
<\psi_{n',\kappa',j',m'_j}|H_I|\psi_{n,\kappa,j,m_j}> = \int_0^\infty dr \int_0^\pi sin\theta d\theta \int_0^{2\pi} d\phi c A(t) r \\
\psi^\dagger_{n',\kappa',j,m_j}
\begin{pmatrix}
iQ_{n,\kappa}(r)\Big(C^{j,l_{-\kappa}}_{m_j-\frac{1}{2},\frac{1}{2}} Y_{l_{-\kappa},m_j-\frac{1}{2}} \begin{pmatrix}1\\0\end{pmatrix} - C^{j,l_{-\kappa}}_{m_j+\frac{1}{2},-\frac{1}{2}} Y_{l_{-\kappa},m_j+\frac{1}{2}} \begin{pmatrix}0\\1\end{pmatrix}\Big) \\
P_{n,\kappa}(r)\Big(C^{j,l_\kappa}_{m_j-\frac{1}{2},\frac{1}{2}} Y_{l_\kappa,m_j-\frac{1}{2}} \begin{pmatrix}1\\0\end{pmatrix} - C^{j,l_\kappa}_{m_j+\frac{1}{2},\frac{1}{2}} Y_{l_\kappa,m_j+\frac{1}{2}} \begin{pmatrix}0\\1\end{pmatrix}\Big)
\end{pmatrix}
\end{aligned}
\tag{3.65}
$$

It is known that spherical harmonics Y_{l,m_l} are orthonormal;

$$
\int_0^\pi \sin\theta d\theta \int_0^{2\pi} Y_{l',m'_l} Y_{l,m_l} = \delta_{l,l'} \delta m_l, m'_l, \tag{3.66}
$$

which transforms (3.65) to

$$
\begin{aligned}
<\psi_{n',\kappa',j',m'_j}|H_I|\psi_{n,\kappa,j,m_j}> = icA(t) \\
\Bigg(\int_0^\infty dr P^*_{n',\kappa'} Q_{n,\kappa} \delta_{l_{-\kappa},l'_\kappa} \delta_{m_j,m'_j} \Big(C^{j',l'^*_\kappa}_{m'_j-\frac{1}{2},\frac{1}{2}} C^{j,l_{-\kappa}}_{m_j-\frac{1}{2},\frac{1}{2}} - C^{j',l'^*_\kappa}_{m'_j+\frac{1}{2},-\frac{1}{2}} C^{j,l_{-\kappa}}_{m_j+\frac{1}{2},-\frac{1}{2}} \Big) \\
+ \int_0^\infty dr Q^*_{n',\kappa'} P_{n,\kappa} \delta_{l_{-\kappa'},l_\kappa} \delta_{m_j,m'_j} \Big(C^{j',l_{-\kappa'}{}^*}_{m'_j-\frac{1}{2},\frac{1}{2}} C^{j,l_\kappa}_{m_j-\frac{1}{2},\frac{1}{2}} - C^{j',l_{-\kappa'}{}^*}_{m'_j+\frac{1}{2},-\frac{1}{2}} C^{j,l_\kappa}_{m_j+\frac{1}{2},-\frac{1}{2}} \Big) \Bigg)
\end{aligned}
\tag{3.67}
$$

We notice that m'_j and m_j must be equal to get non-zero results, which shows that the dipole approximation yields decoupled results for different m_j. We previously stated that our initial state has $m_j = \frac{1}{2}$ and therefore $m'_j = \frac{1}{2}$. We write[37]

$$
\begin{aligned}
<\psi_{n',\kappa',j',\frac{1}{2}}|H_I|\psi_{n,\kappa,j,\frac{1}{2}}> = icA(t) \\
\Bigg(\int_0^\infty dr P^*_{n',\kappa'} Q_{n,\kappa} \delta_{l_{-\kappa},l'_\kappa} \Big(|C^{j,l_{-\kappa}}_{0,\frac{1}{2}}|^2 - |C^{j,l_{-\kappa}}_{1,-\frac{1}{2}}|^2 \Big) \\
+ \int_0^\infty dr Q^*_{n',\kappa'} P_{n,\kappa} \delta_{l_{-\kappa'},l_\kappa} \Big(|C^{j,l_\kappa}_{0,\frac{1}{2}}|^2 - |C^{j,l_\kappa}_{1,-\frac{1}{2}}|^2 \Big) \Bigg) \\
= icA(t) \Bigg(\Big(\delta_{\kappa,-\kappa'} + \delta_{\kappa,\kappa'+1} \Big) \int_0^\infty dr P^*_{n',\kappa'} Q_{n,\kappa} \Big(|C^{j,l_{-\kappa}}_{0,\frac{1}{2}}|^2 - |C^{j,l_{-\kappa}}_{1,-\frac{1}{2}}|^2 \Big) \\
+ \Big(\delta_{\kappa,-\kappa'} + \delta_{\kappa,\kappa'-1} \Big) \int_0^\infty dr Q^*_{n',\kappa'} P_{n,\kappa} \Big(|C^{j,l_\kappa}_{0,\frac{1}{2}}|^2 - |C^{j,l_\kappa}_{1,-\frac{1}{2}}|^2 \Big) \Bigg),
\end{aligned}
\tag{3.68}
$$

where the radial integrals are numerically determined. Obviously we cannot work with $r \in [0,\infty)$, as it requires infinite numerical power. We choose to truncate r at R_{max}, such that $r \in [0, R_{max}]$.

The next step is to define the vector field $A(t)$. We want the vector field $A(t)$ to mimic a realistic

physical pulse, and therefore it is developed as a cosine function inside an envelope function

$$A(t) = \frac{E_0}{\omega}\left(\sin\left(\frac{\pi t}{T_{pulse}}\right)\right)^2 \cos\left(\omega t + \phi\right), \tag{3.69}$$

where T_{pulse} is the duration of the pulse, ω is the angular frequency, ϕ is the phase of the pulse and E_0 is the amplitude of the electric field strength. We will choose to look at a pulse which has 15 periods and therefore $T_{pulse} = \frac{30\pi}{\omega}$. The vector field will have the form

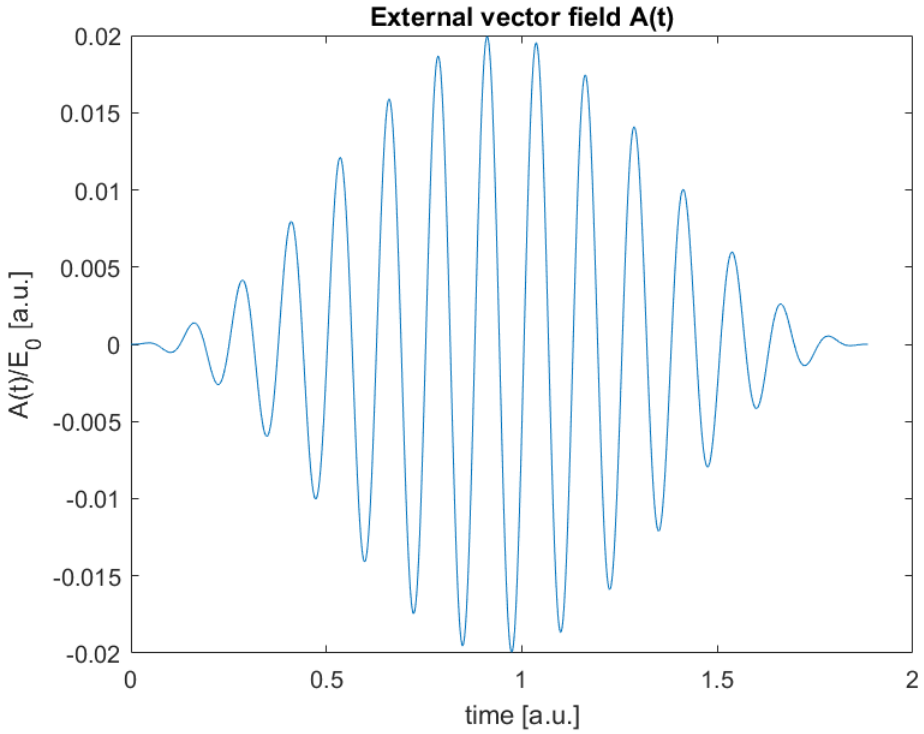


Figure 3.14: Illustration of vector field A(t) with $\omega = 50$, $T_{pulse} = \frac{30\pi}{\omega}$ and $\phi = -\frac{\pi}{2}$.

3.10 Krylov Subspace Methods

When solving a non-singular linear system of the form[25]

$$Ax = b \tag{3.70}$$

it is normal to use Gaussian elimination. Although that can be a numerical challenge when the matrix A is of a significant size. It therefore calls for an approximating method which can decrease the size of the system that is to be solved. Such methods can be Krylov subspace methods. The Krylov subspace methods are iterative processes, where one wants to solve the system

in (3.70) for x. If one chooses to do k iterations, the approximated solution x_k will be a linear combination of the basis vectors in the Krylov space $\mathcal{K}_k$,

$$\mathcal{K}_k = span\big(c, Ac, ...A^{k-1}c\big), \tag{3.71}$$

where c is an arbitrary vector that needs to be chosen. A popular choice is $c = b$, which we also will do in this thesis. One searches naturally for x_k which solves the least squares problem

$$\min_{x_k \epsilon \mathcal{K}_k(A,b)} \quad ||b - Ax|| \tag{3.72}$$

It can be solved by constructing an orthonormal basis $\{v_1, v_2, ...v_k\}$ for the Krylov space $\mathcal{K}_k$. We construct the orthonormal basis by using Arnoldi's method, which is a modified Gram-Schmidt procedure. The Gram-Schmidt procedure constructs orthonormal vectors $\{v_1, v_2, ...v_k\}$ for the Krylov space $\mathcal{K}_k = span(b, Ab, ..A^{k-1}b$ in this manner;

$$\begin{aligned}
v_1 &= \frac{b}{||b||} \\
q_2 &= Ab - (b^* Ab)\, Ab \\
v_2 &= \frac{q_2}{||q_2||} \\
&\vdots \\
q_k &= A^{k-1}b - \sum_{j=1}^{k-1} (v_j^* A^{k-1})\, v_j \\
v_k &= \frac{q_k}{||q_k||}
\end{aligned} \tag{3.73}$$

It is quite numerical expensive to do all these calculations. It is advantageous to use the Arnoldi's method instead. In the Gram-Schmidt procedure one finds the orthonormalized vector v_j by projecting $A^{j-1}b$ against the previously orthonormalized vectors $\{v_1, .., v_{j-1}\}$. The idea behind Arnoldi's method is to find v_j by projecting Av_{j-1} against the previously orthonormalized vec-

tors $\{v_1,..,v_{j-1}\}$. The Krylov subspace $\mathcal{K}_k$ can be rewritten as[25]

$$\begin{aligned}
\mathcal{K}_k &= span(b, Ab, A^2b...., A^{k-1}b) \\
&= span(v_1, Av_1, A^2v_1..., A^{k-1}v_1) && \text{as } v_1 = \frac{b}{||b||} \\
&= span\Big(v_1, \alpha v_1 + \beta v_2, A(\alpha v_1 + \beta v_2), ..., A^{k-2}(\alpha v_1 + \beta v_2)\Big) && \text{as } Av_1 = \alpha v_1 + \beta v_2 \\
&= span\big(v_1, v_2, Av_2, ..., A^{k-2}v_2\big) \\
&\quad \vdots \\
&= span\big(v_1, v_2, ..., Av_{k-1}\big)
\end{aligned} \tag{3.74}$$

Therefore the j-th orthonormal basis vetor v_j can be written as;

$$\begin{aligned}
q_j &= Av_{j-1} - \sum_{i=1}^{j-1} (v_i^* Av_{j-1})\, v_i \\
v_j &= \frac{q_j}{||q_j||}
\end{aligned} \tag{3.75}$$

v_j and q_j is obviously parallel and therefore

$$\begin{aligned}
v_j * q_j &= ||q_j|| \\
&\overset{(3.75)}{=} v_j^* Av_{j-1}
\end{aligned} \tag{3.76}$$

If we let

$$m_{ij} = v_i^* Av_j \tag{3.77}$$

We can make use of the relations from (3.76) and (3.75) and write

$$Av_j = \sum_{i=1}^{j+1} v_i m_{ij} \tag{3.78}$$

It would be advantageous for numerical purposes to get a matrix representation. We therefore define

$$V_k = \big[v_1, v_2, ..., v_k\big], \tag{3.79}$$

which leads to the Arnoldi relation

$$AV_k = V_k M_k + \Big[\underbrace{0, ..., 0}_{k-1 \text{ times}}, v_{k+1} m_{k+1,k}\Big] \tag{3.80}$$

If one chooses k such that $m_{k+1,k} = 0$, we have constructed a basis that span the Krylov space $\mathscr{K}_k$. Otherwise one makes the approximation

$$AV_k \approx V_k M_k \tag{3.81}$$

If the matrix A is hermitian, the Arnoldi procedure can be further simplified to the Lanczos procedure. It can be shown that[25]

$$AV_k \approx V_k T_k, \tag{3.82}$$

where T_k is a tridiagonal matrix. As V_k is a matrix where the columns are orthonormal, it must obviously be unitary;

$$V_k^\dagger V_k = \mathbb{1}_k, \tag{3.83}$$

which gives

$$A \approx V_k T_k V_k^\dagger \tag{3.84}$$

The Krylov subspace method will show itself to be very useful for solving the Dirac equation with an external pulse. If one let $A = H(t)$, it will drastically decrease the number of numerical operations necessary for developing our hydrogen atom in time. See the next section for more details.

3.11 Application of Krylov methods

In section 3.8 we were given the equation

$$\begin{aligned} c(t+\Delta t) &= \left(\sum_{n=0}^{\infty} \frac{\big(-iH(t)\Delta t\big)^n}{n!} \right) c(t) \\ &= \sum_{n=0}^{\infty} \frac{\big(-i\Delta t\big)^n}{n!} H^n(t) c(t) \end{aligned} \tag{3.85}$$

Let us now create a Krylov subspace $\mathscr{K}_k\big(H(t), c(t)\big)$. As $H(t)$ is obviously hermitian, the Lanczos procedure can be useful. It lets us write $H(t)^n$ as[39]

$$H(t)^n = V T^n V^\dagger \approx V_k T^n V_k^\dagger, \tag{3.86}$$

where the columns of V are the orthonormal basis vectors for the Krylov space $\mathscr{K}_k$ created by the Lancszos algorithm and T is a tridiagonal matrix. We insert it into (3.53):

$$c(t+\Delta t) \approx \sum_{n=0}^{\infty} \frac{\big(-i\Delta t\big)^n}{n!} V_k T^n V_k^\dagger c(t) \tag{3.87}$$

As $c(t)$ is the first column of $V_k \Rightarrow V_k^\dagger c(t) = e_1$, which simplifies (3.53) further to

$$\begin{aligned} c(t+\Delta t) &\approx V_k \sum_{n=0}^{\infty} \frac{\big(-i\Delta t\big)^n}{n!} T^n e_1 \\ c(t+\Delta t) &\approx V_k \sum_{n=0}^{L} \frac{\big(-i\Delta t\big)^n}{n!} T^n e_1, \end{aligned} \tag{3.88}$$

where we have at the last line have truncated the series at $n = L$. We emphasize that V_k need to be calculated at every time step, by using the Lanczos procedure. Before we applied the Krylov subspace method, we needed to calculated $H^n(t)$ in (3.85). This is a severe numerical process, as $H(t)$ is of a significant size. The dimension of the tridiagonal matrix T in (3.88) is $k \times k$, which can be chosen to be significantly less than the dimension of $H(t)$. In that regard, the Krylov subspace methods save us an immensely amount of numerical operations and time.

When (3.88) is numerically solved, we are interested in investigating photoionzation of an electron in the hydrogen ground state to the positive pseudo continuum. Techniques for such investigations are discussed in the next section.

3.12 Investigation of Photoionization

It can be interesting to investigate the probability of an electron in the ground state to be excited to positive pseudo continuum states when it interacts with an external pulse. We know from basic quantum mechanics that the probability P_i to find an electron in the $i-th$ state at a time t' is given by[27]

$$P_i = c_i(t')^* c_i(t') = |c_i(t)|^2, \tag{3.89}$$

where $c(t)$ is the time-dependent coefficient from equation (3.88) in the previous section. As we want to model the hydrogen ground state as the initial state at $t = 0$, we choose

$$c_{gs}(0) = 1 \tag{3.90}$$

where c_{gs} is the coefficient for the ground state. We want to investigate the probability of finding the electron in the positive pseudo continuum, so we must naturally find the coefficients $c_{pos.cont.}$ for the states in the positive pseudo continuum. This is done by solving (3.88) numerically. We remember that the positive pseudo continuum is, as its name suggest, a quantized continuum. We would ideally want to model a physical positive continuum, so we can define a

property that is relatable to the physical continuum, namely the differential probability

$$\frac{dP_i}{dE} = P_i\sigma(E_i), \tag{3.91}$$

where $\sigma(E_i)$ is the energy density for the $i-th$ state, and it can be approximated as

$$\sigma(E_i) \approx \frac{2}{E_{i+1} - E_{i-1}} \tag{3.92}$$

The differential probability will give us an indication of how the photoionzation would behave if we indeed had a continuous positive continuum.

Chapter 4

Results

We have in the previously chapters laid the groundwork for what is up next, namely the results. We follow the same path of logic as earlier, as we solve the radial Dirac equation without an external pulse first and secondly solve the time-dependent Dirac equation with an external pulse. This order is necessary, because the results for the radial Dirac equation without an external pulse will be used to solve the time- dependent Dirac equation with an external pulse.

4.1 Solutions to the Radial Dirac Equation Without an External Pulse

When solving the radial Dirac equation numerically, it is of great importance to be able to differ between physical results and results that appear due to numerical limitations. In that regard, analytic results are useful as they can be compared with numerical results. We will firstly identify relativistic effects by comparing relativistic results, both numerical and theoretical, from the the radial Dirac equation with results from the NR, non-relativistic, Schrödinger equation. Secondly, we will look at convergence properties and numerical limitations when choosing the size of our basis set $2N$ and the size of the radial interval $r \in [0, R_{max}]$.

4.1.1 Identification of Relativistic Effects

Firstly, we will look at various solutions to the radial Dirac equation without an external pulse. We will compare both radial wavefunctions ψ_{rad} and energies E solutions yielded from the radial Dirac equation with corresponding solutions to the NR Schrödinger equation. In this regard, the goal is to identify relativistic effects.

The relativistic effects observed are obviously dependent on the nucleus charge Z, as the electron's mean velocity increases as Z increases. Therefore one expects to observe more distinct relativistic effects for higher nucleus charge Z. We start by look at hydrogen with $Z = 1$, where one does not expect very distinct relativistic effects. Some relativistic effects can though be observed in figure 4.1.

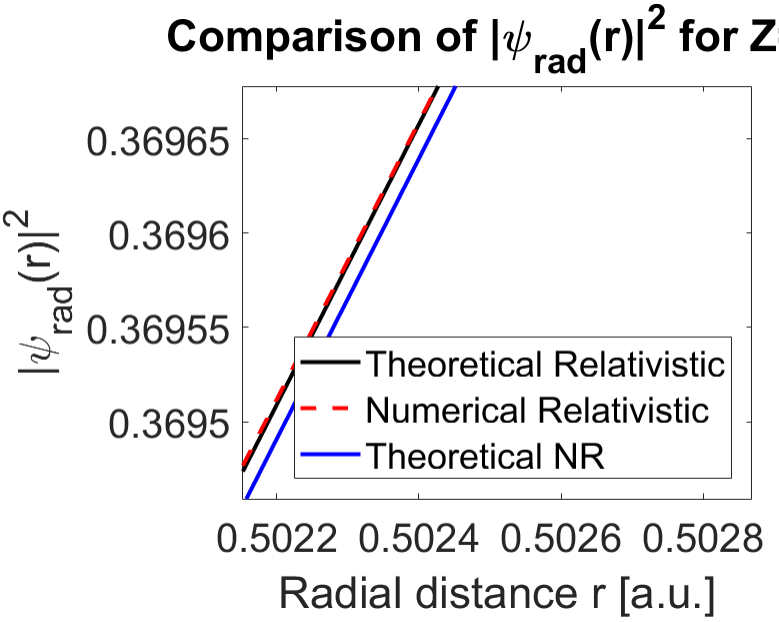


Figure 4.1: Comparison of numerical relativistic-, theoretical relativistic- and theoretical NR radial probability $|\psi_{rad}|^2$ for hydrogen where $Z = 1$. We note that the numerical computations are with $N = 400$.

Figure 4.1 shows, as one expects, that the electron will have higher probability, both numerical and theoretical, to exist closer to the nucleus in the relativistic Dirac equation than in the NR Schrödinger equation. This is a relativistic effect, but it is quite small, as the shift is in the order of 10^{-5}. This relativistic effect is due to length contraction and is expected to be more distinct for higher Z. It is also noteworthy that figure 4.1 shows great compliance between the theoretical and numerical relativistic radial probability $|\psi_{rad}|^2$, which indicates that the numerical radial probability is a good approximation. One expected more distinct length contraction for larger Z, and this is observed for $Z = 40$ in figure 4.2a, $Z = 50$ in 4.2b,..., $Z = 80$ in figure 4.2e and $Z = 92$ in figure 4.2f.

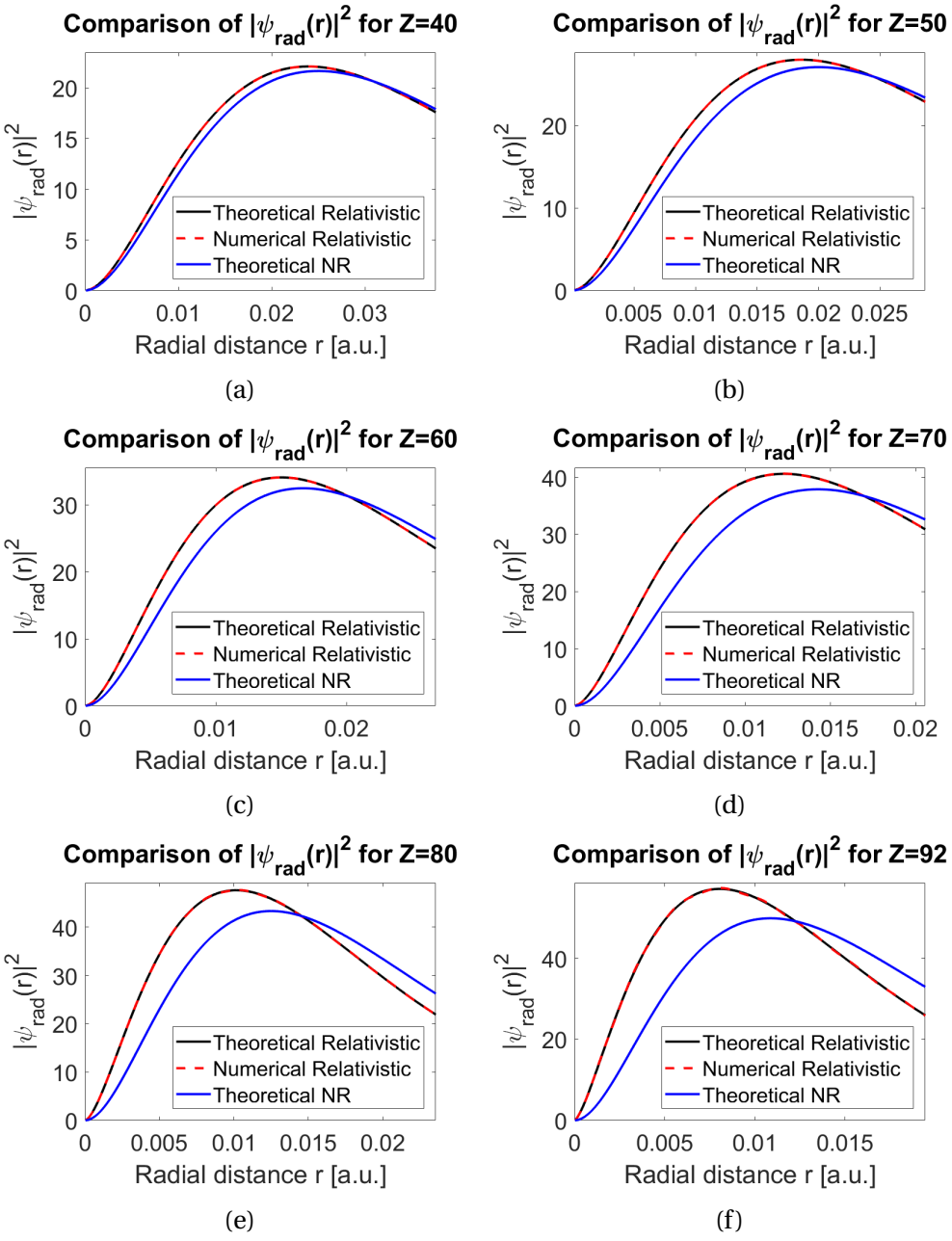


Figure 4.2: Comparison of numerical relativistic-, theoretical relativistic- and theoretical NR radial probability $|\psi_{rad}|^2$ for (a) $Z = 40$ (b) $Z = 50$ (c) $Z = 60$ (d) $Z = 70$ (e) $Z = 80$ (f) $Z = 92$. It is for a hydrogenic ground state and all numerical computations are with $N = 400$.

We observe in the figure above that higher nucleus charge Z indeed causes more distinct relativistic effects. The difference between the NR and relativistic radial probabilities increases from $Z = 40$ to $Z = 92$, which states that the electron will have higher probability of existing near the nucleus, compared to its corresponding NR, non-relativistic, treatment. This is caused by length contraction of the radial distance within the atom. Length contraction of the atom

will have other physical consequences as well. The electron will have higher probability of being close to the nucleus, which let us expect that the Coulomb potential will become greater in absolute value. With that in mind, it is expected that the energies of the electron will be lower in the relativistic description than in the NR description. This is verified by the results in table 4.1.

Table 4.1: Comparison of ground state energy for (i) numerical relativistic treatment (ii) theoretical relavistic treatment (iii) theoretical NR treatment of the electron in ground state. The nucleus charge Z is varied and we note that $N = 400$ in the numerical treatment.

Nucleus charge\Energies	**Numerical rel. [a.u.]**	**Theoretical rel. [a.u.]**	**Theoretical NR [a.u.]**
Z=1	–0.500006656598998	-0.500006656598998	-0.500000000000000
Z=40	-817.8082	-817.8074	-800.0000
Z=50	-1294.6290	-1294.6261	-1250.0000
Z=60	-1895.690	-1895.682	-1800.000
Z=70	-2624.868	-2634.847	-2450.000
Z=80	-3532.24	-3432.19	-3200.00
Z=92	-4861.48	-4861.20	-4232.00

Table 4.1 states that the difference between the relativistic ground state energies and the NR ground state energies increases as Z increases. It is also noteworthy that the difference between the numerical relativistic ground state energy and the theoretical ground state energy also increases as Z increases. They are equal up the order of 10^{-16} for $Z = 1$, but unequal in the first decimal for $Z = 92$. This stems from the fact that the radial probability $|\psi_{rad}|^2$ increases for small r as Z increases, which yields numerical difficulties in handling the divergence of $\frac{1}{r}$-terms as r approaches zero in the radial Dirac equation.

4.1.2 Convergence Properties of the Radial Dirac Equation

We were in chapter 3 introduced to certain truncation factors which were applied to make the radial Dirac equation numerical solvable. Such truncation factors were the number of basis functions $2N$ and R_{max}, the truncation of the r-grid. $2N$ number of basis functions yield $2N$ solutions in the radial Dirac equation, and it therefore determines the range of energies in both the positive- and negative energy pseudo continuum. We are going to inspect photoionization of hydrogen in the next section, where the energies in the positive pseudo continuum are going to be of great importance. In that regard, $2N$ will determine which positive pseudo continuum states the electron can be excited to.

The number of basis functions $2N$ will not greatly influence the number of bound states present

in the solution of the radial Dirac equation, but we need it to be above a certain threshold, as we want to be able to mimic the behaviour of physical bound states. The limited amount of influence $2N$ has on bound states can be observed in table 4.2, table 4.3 and table 4.4. The absolute difference between the theoretical- and numerical energies are observed to be nearly independent of $2N$.

The choice of R_{max} has, on the other hand, great influence on the bound states. The smaller R_{max} is, the more curvature will ψ_{rad} have, which is an indication of higher kinetic energy. Also, according to Heisenberg's uncertainty relation, will the uncertainty in the kinetic energy increase if we decrease R_{max}. In sum will a decrease in R_{max} decrease the numerical precision for the bound excited states and also decrease the amount of bound states solutions. The difference in numerical precision for bound excited states as we vary R_{max} is shown for $R_{max} = 125$ in table 4.2, $R_{max} = 100$ table 4.3 and $R_{max} = 60$ table 4.4. We emphasize that n indeed is the main principle quantum number in the three tables below.

Table 4.2: Comparison of absolute difference between theoretical and numerically solved energies $E_{n,\kappa=-1}$ for $n = 1,..6$ for different number of basis functions $2N$ with $R_{max} = 125$. The least precise numerical results are marked with red background.

N\\|$E_{num} - E_{theory}$|	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
100	3.6380e-12	7.2760e-12	7.2760e-12	1.0914e-11	2.5466e-11	1.4572-07
200	7.2760e-12	7.2760e-12	1.0914e-11	0	1.8190e-11	1.4572e-07
300	3.6380e-12	0	0	3.6380e-12	1.4552e-11	1.4572e-07
400	7.2760e-12	3.6380e-12	1.8190e-11	3.6380e-12	1.4552e-11	1.4571e-07
500	0	0	1.4552e-11	3.6380e-12	1.0914e-11	1.4571e-07

Table 4.3: Comparison of absolute difference between theoretical and numerically solved energies $E_{n,\kappa=-1}$ for $n = 1,..6$ for different number of basis functions $2N$ with $R_{max} = 100$. The least precise numerical results are marked with red background.

N\\|$E_{num} - E_{theory}$|	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
100	1.0914e-11	1.8190e-11	3.6380e-12	3.6380e-12	2.8496e-08	2.0409e-05
200	0.0000	3.6380e-12	3.6380e-12	3.6380e-12	2.8493e-08	2.0409e-05
300	0.0000	0.0000	1.0914e-11	3.6380e-12	2-8500e-08	2.0409e-05
400	0.0000	3.6380e-12	0.0000	3.6380e-12	2.8489e-08	2.0409e-05
500	7.2760e-12	3.6380e-12	1.0914e-11	0.0000	2.8493e-08	2.0409e-05

Table 4.4: Comparison of absolute difference between theoretical and numerically solved energies $E_{n,\kappa=-1}$ for $n = 1,..6$ for different number of basis functions $2N$ with $R_{max} = 60$. The least precise numerical results are marked with red background.

N\\$\lvert E_{num} - E_{theory}\rvert$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
100	1.4552e-11	0	4.0018e-11	1.8500e.06	4.0347e-05	0.0044
200	0	1.4552e-11	2.9104e-11	1.8500e-06	4.0347e-05	0.0044
300	0	3.6380e-12	2.9104e-11	1.8500e-06	4.0347e-05	0.0044
400	3.6380e-12	7.2760e-12	3.2742e-11	1.8500e-06	4.0347e-04	0.0044
500	7.2760e-12	3.6380e-12	2.9104e-11	1.8500e-06	4.0347e-04	0.0044

We observe in table 4.2, table 4.3 and table 4.4 that when R_{max} decreases, so does the numerical precision for the bound excited states. $R_{max} = 125$ in table 4.2 yields sufficient precision up to fourth excited state ($n = 5$), but is imprecise for the fifth excited state, marked in red. $R_{max} = 100$ yields, on the other hand, sufficient precision up to the third excited state and insufficient precision for the fourth and fifth excited state. Finally, $R_{max} = 60$ yields only sufficient precision up to the second excited state and is imprecise for the third, fourth and fifth excited state, also marked in red.

We will in the next section look at solutions to the time-dependent Dirac equation with an external pulse. We will look at the hydrogen ground state as the initial state and investigate excitation to states in the positive pseudo continuum. In that regard will bound excited states be of negligible importance and $R_{max} = 60$ will be sufficient, as it yields precise results for the ground state energy.

4.2 Solutions to the Time-Dependent Dirac Equation with an External Pulse

We will in this section solve the time-dependent Dirac equation with an external pulse. The external pulse is within the dipole approximation, as the spatial part is neglected. We select a pulse of the form

$$A(t) = \frac{E_0}{\omega}\left(\sin\left(\frac{\pi t}{T_{pulse}}\right)\right)^2 \cos\left(\omega t + \phi\right), \tag{4.1}$$

and we choose $\omega = 50$, $T_{pulse} = 15\pi$ and $\phi = -\frac{\pi}{2}$. It is the same pulse that was shown in figure 3.14 in chapter 3, section 3.9.

We want to investigate, as stated earlier, photoionization of a ground state electron in hydro-

gen to positive pseudo continuum states. In that regard, one must choose the electron in ground state as the initial state and then develop it in time by using Krylov subspace methods (see chapter 3, section 3.11).

4.2.1 Convergence Properties of the Time-Dependent Dirac Equation with an External Pulse

We will also meet upon truncation factors for the time-dependent Dirac equation with an external pulse, as we did for the radial Dirac equation without an external pulse. Such truncation factors are the size of the Krylov subspace k, the timestep Δt and crucially the number of κ-values we include in our numerical computations. We will not look deeply into the convergence properties of k and Δt, although choosing $\Delta t = 5000$ and $k = 6$ has shown itself to be sufficient.

When we looked at the convergence properties for the radial Dirac equation without an external pulse in section 4.1.2, we emphasized that the number of basis functions $2N$ chosen determines the range of energies yielded in the solutions. Specifically, it determines the maximum energy in the positive pseudo continuum. This will be of importance when investigating photoionization of the electron in the ground state, as it may be excited multiple times. The electron will have a probability, up to a certain order, to be excited to high energy states in the positive pseudo continuum. With that in mind, it is important to choose a sufficient amount of basis functions $2N$, so that these high energy states may be represented in our numerical solution.

The number of κ-values we include in our computations is also crucial. Ideally, we would to include all κ-values, but that is a numerical impossibility. The number of κ does, in many ways, determine the density of states in the positive pseudo continuum and it will influence the differential probability energy distribution, as seen in figure 4.3.

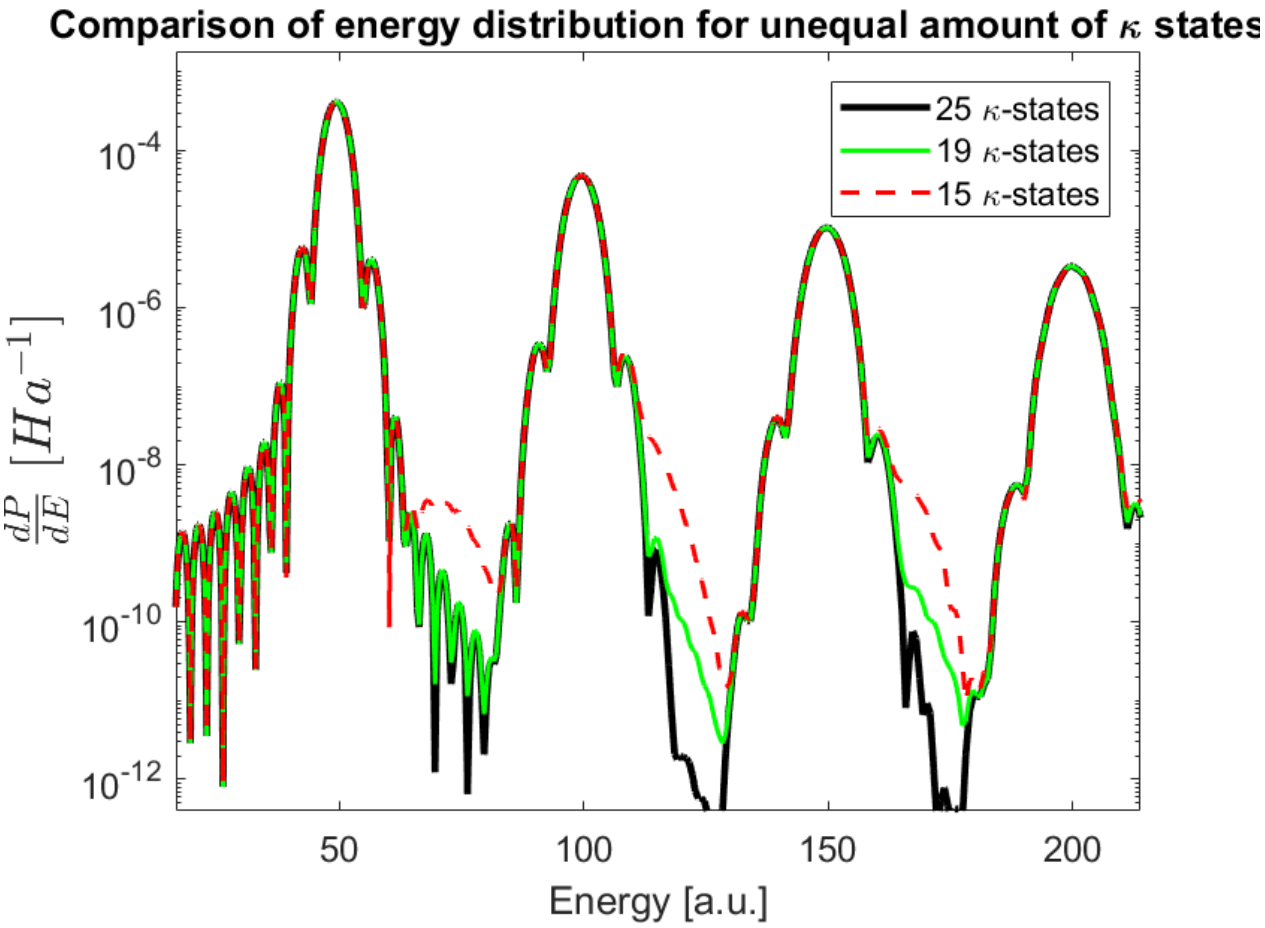


Figure 4.3: Comparison of differential probability energy distributions for computations with a different amount of κ-states. The black line with 25 κ-states is with -13 $\leq \kappa \leq -1$ and $1 \leq \kappa \leq 12$, the green line with 19 κ-states is with -10 $\leq \kappa \leq -1$ and $1 \leq \kappa \leq 9$, and the red dotted line with 15 κ-states is with -8 $\leq \kappa \leq -1$ and $1 \leq \kappa \leq 7$. We note that the results are computed with R_{max}=40 and $E_0 = 600$.

We interpret figure 4.3 as a probability distribution over energies. With that in mind, we can interpret the first peak as the probability of an electron being excited once and the second peak as the probability being excited twice, and so on. If one increases the number of basis functions $2N$, one would be able to observe even more peaks. We note that we would ideally like to get all the possible peaks into our computations, but that would require $2N$ to be infinite. Instead, we must settle for a decided precision. One could also observe in figure 4.3 that the differential probability at the peaks decreases as the energy decreases, so one expects the next physical peak, which is not in our calculations, to have lower differential probability than the last computed peak. As such, one could expect the next not computed peak in figure 4.3 to be in the order of 10^{-6}.

It is also observed in figure 4.3 that the amount of κ-values included in the computations will affect the differential probability energy distribution. The largest difference for computations with different amount of κ-values is found between the peaks, but it is only in the order of 10^{-7}. That is smaller than the expected differential probability for the next peak that is not computed, so in this case will the difference be negligible. We note that increasing both $2N$ and the number of κ-values will greatly increase the computational time, and that an increase in $2N$ would require more κ-states. This is seen in figure 4.4 where a greater number of κ-values is needed for full convergence.

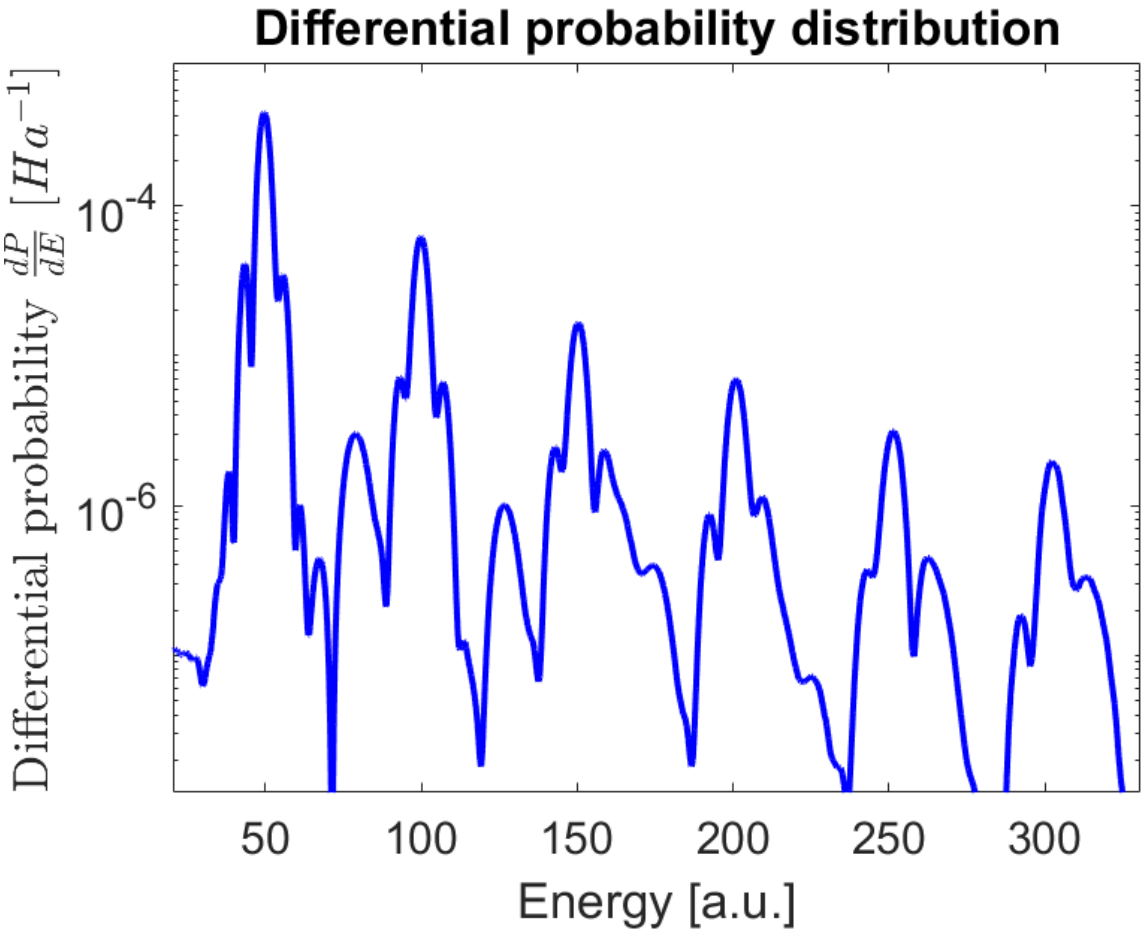


Figure 4.4: Not fully converged differential probability energy distribution computed with $E_0 = 1000$, $N = 520$, $R_{max} = 60$ and 11 κ-states.

We observe that the differential probability in figure 4.4 is quite unsymmetrical and not as smooth as we observed for the computation with 25 κ-states in figure 4.3. This indicates that the differential probability is not fully converged, but it does not seem like the precision for the total prob-

ability for photoionzation to positive pseudo continuum states is too affected. It is observed in figure 4.4 that 11 κ-states for $E_0 = 1000$ yield a precision in the region between 10^{-5} and 10^{-6} for the total probability of photoionization and it is assumed to be sufficient for further computations in the next section.

4.2.2 Identifying Relativistic Effects

As the convergence properties of the time-dependent Dirac equation with en external pulse is established, it is time to compare the relativistic results with results from the NR Schrödinger. We do this in order to identify relativistic effects in our computations. We are still looking at photoionization for an electron in the hydrogen ground state to positive pseudo continuum states. Firstly, let us compare the differential probability energy distribution for the time-dependent Dirac equation with a corresponding distribution for the NR Schrödinger. We note that the data for the NR differential probability energy distribution is received from the author of[40].

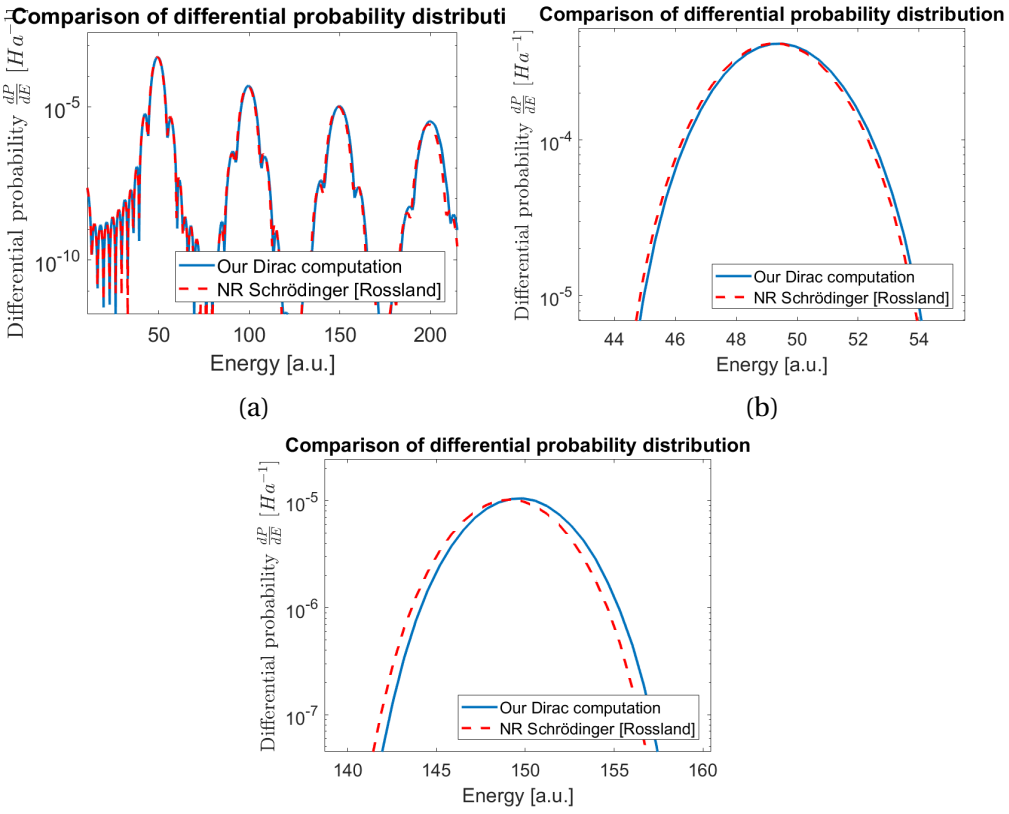


Figure 4.5: Comparison of differential probability distribution for our numerical Dirac solution and received solutions of the NR Schrödinger equation from[40]. Figure (a) shows the full distribution, figure (b) is the same figure zoomed in on the first peak, and figure (c) is zoomed in on the third peak. We note that all numerical computations are made with $E_0 = 600$ and $R_{max} = 60$. The calculations for Dirac are also computed with 11 κ-states and $N = 300$.

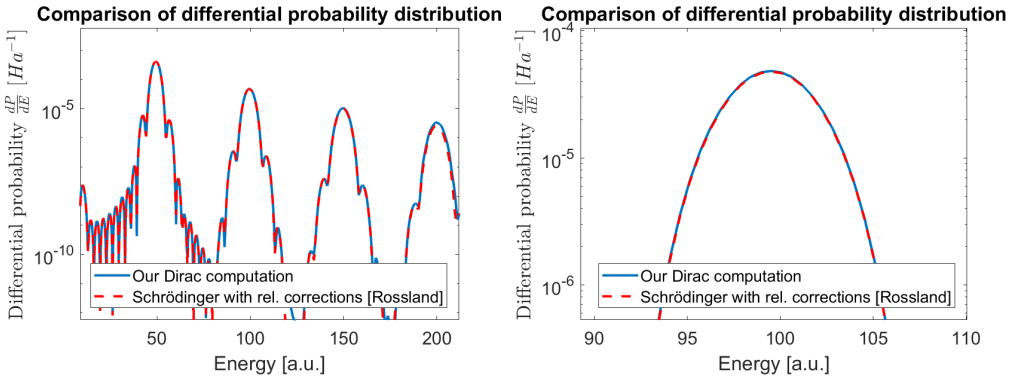
Figure 4.5a shows that the differential probability distribution for the Dirac equation and the NR Schrödinger equation[40] matches quite well. This is to be expected as $E_0 = 600$ is not a too strong electric field strength, but some effects can actually be observed if one zoom in on the different peaks. We have in figure 4.5b zoomed into the first peak, where one can observe that the NR peak from the Schrödinger equation is shifted to the left compared to the corresponding peak from the Dirac equation. The shift in figure 4.5b also seem to accumulate for higher energy peaks, observed in figure 4.5c.

It can be convenient to check whether these shifts disappear for a Schrödinger equation with relativistic corrections. In that regard, there has been proposed a Schrödinger equation with

relativistic corrections, where the Hamiltonian $\hat{H}^S_{rel}$ is given by[40]

$$\hat{H}^S_{rel} = -\frac{\hat{p}^2}{2} - \frac{1}{r} + \hat{p}_z\mathbf{A} - \frac{\hat{p}^4}{8c^2} - \frac{\hat{p}^2\hat{p}_z\mathbf{A}}{c^2} - \frac{1}{4}\frac{\hat{p}^2\mathbf{A}^2}{c^2} - \frac{(\hat{p}_z\cdot\mathbf{A})(\hat{p}_z\cdot\mathbf{A})}{2c^2} - \frac{\hat{p}_z\mathbf{A}^3}{2c^2} \tag{4.2}$$

The differential probability energy distribution for the Schrödinger equation with relativistic corrections has been solved by[40], and the solutions are shown here for comparison with the differential probability for the Dirac equation. The comparison is shown in both figure 4.6a and figure 4.6b.



(a) (b)

Figure 4.6: Comparison of differential probability distribution for our numerical Dirac solution and received solutions of the Schrödinger equation with relativistic corrections from[40]. Figure (a) shows the full distribution, and figure (b) is the same figure zoomed in on the first peak. We note that all numerical computations are made with $E_0 = 600$ and $R_{max} = 60$. The calculations for Dirac are also computed with 11 κ-states and $N = 300$.

It is observed in figure 4.6a and figure 4.6b that the differential probability for the Schrödinger equation with relativistic corrections[40] matches with the one computed from the Dirac equation. It verifies that the shift seen in figure4.5b indeed is a relativistic effect.

It could also be interesting to investigate relativistic effects in the total probability for photoionization, the transition amplitude, of ground state electron to positive pseudo continuum states. In that regard it could be interesting to investigate the transition amplitude as one varies E_0. One expect to observe increasing relativistic effects as E_0 increases, because the maximum quiver velocity of a free NR electron in a oscillating homogeneous electric field with amplitude E_0 and angular frequency ω is[21]

$$v = \frac{E_0}{\omega} \tag{4.3}$$

We therefore propose to sum all probabilities for the electron to be in the pseudo continuum, as the sum will pose as the total probability for photoionzation of the ground state electron. Numerical solutions for photoionzation have already been proposed for the Dirac equation by[41], shown in figure 4.7.

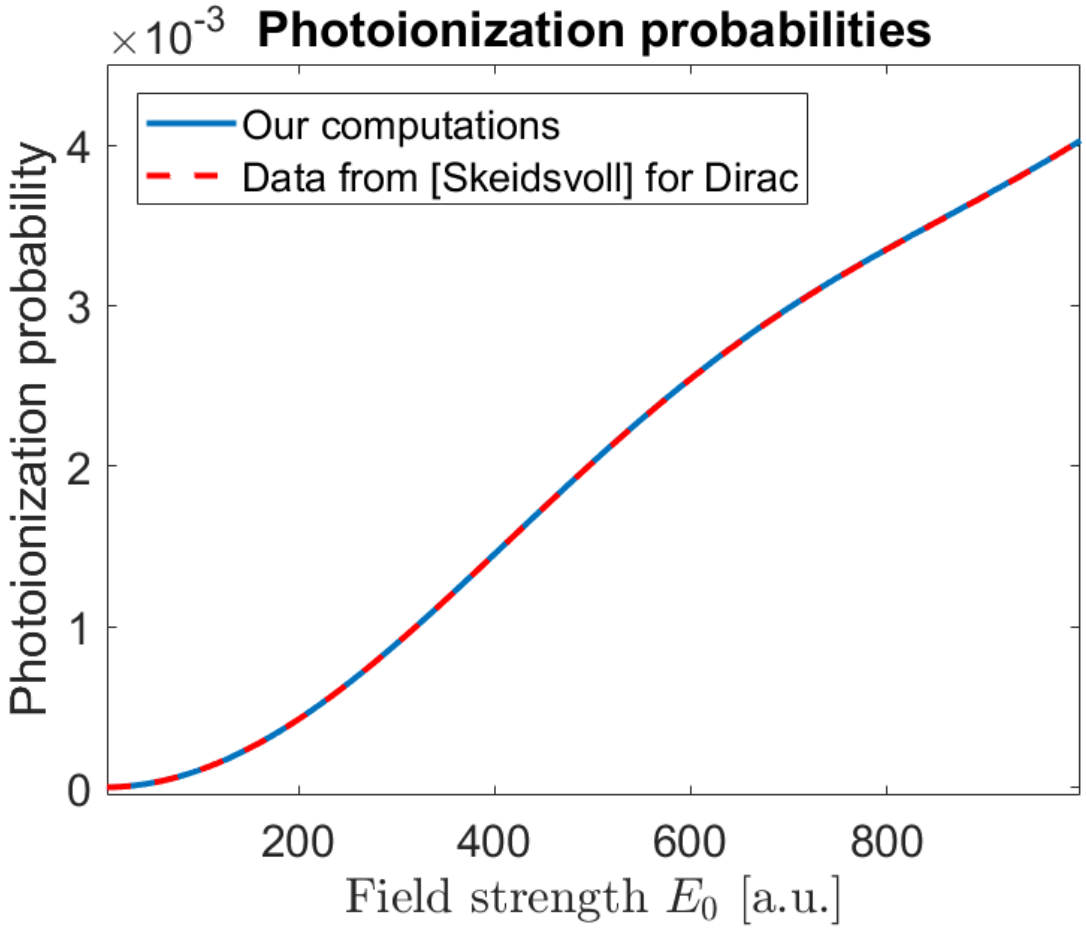


Figure 4.7: Comparison of probabilities for photoionization in the Dirac equation between solutions by[41] and our numerical results for different field strengths E_0.

Figure 4.7 shows that our computed probabilities for photoionization greatly matches the computations in[41]. This is a benchmark, and it verifies the validity of our computations. The figure also shows, as expected, that increased field strength E_0 increases the probability of photoionization.

The next natural step is to determine whether there are any relativistic effects in the probabilities of photoionization. In that case, one must compare the calculated probabilities for photoionization with the corresponding probabilities yielded from solving the same problem in the NR Schrödinger equation. It has been solved by[40] and the probabilities are shown in figure 4.8.

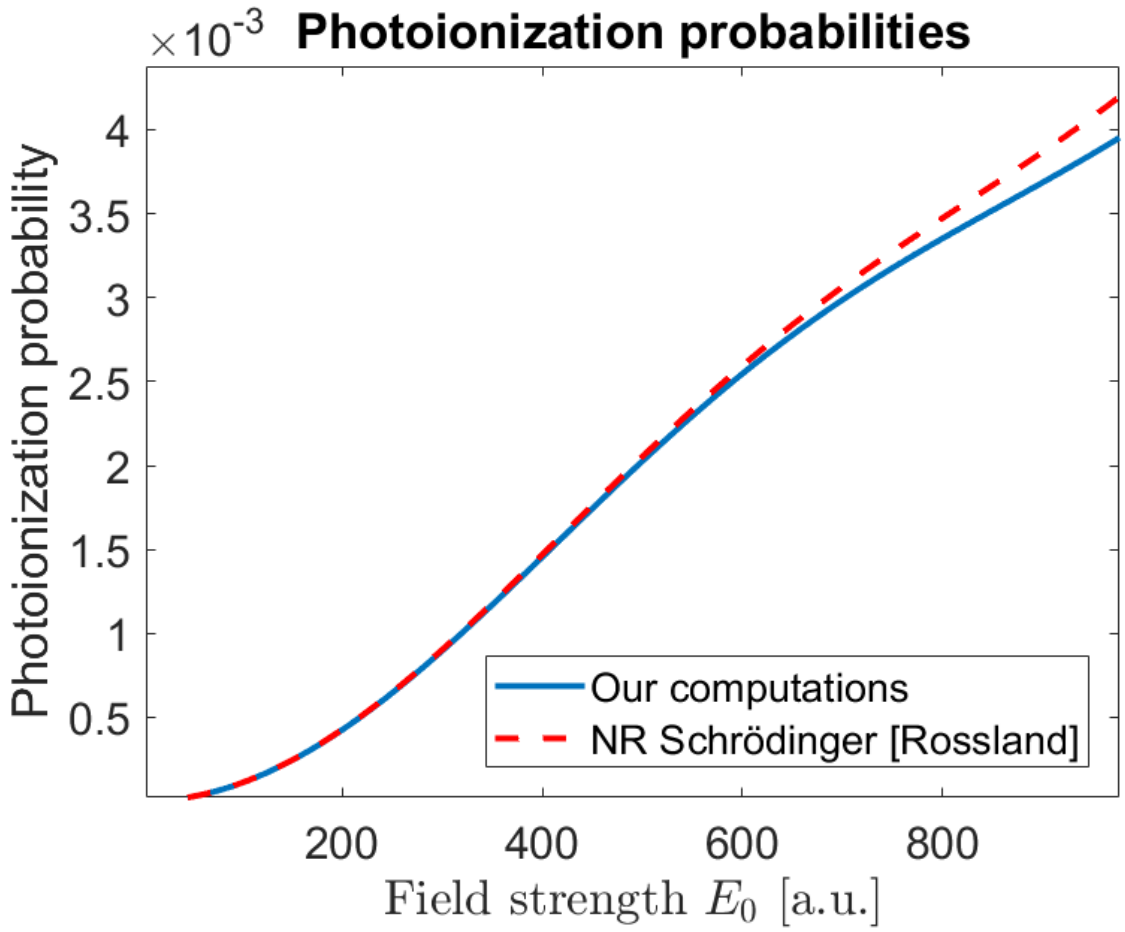


Figure 4.8: Comparison of probabilities for photoionization in the Dirac equation and in the NR Schrödinger equation[40] for different field strengths E_0.

We observe in figure 4.8 that the Dirac equation yields less probability for photoionization than the NR Schrödinger, especially for large field strengths E_0. This is a relativistic effect.

It can also be shown that the Schrödinger equation with relativistic corrections yields consistent probabilities for photoionzation with the Dirac equation.[40] has proposed such probabilities, by using the same Hamiltonian $\hat{H}^S_{rel}$ given in equation (4.2). The probabilities are shown in figure 4.9.

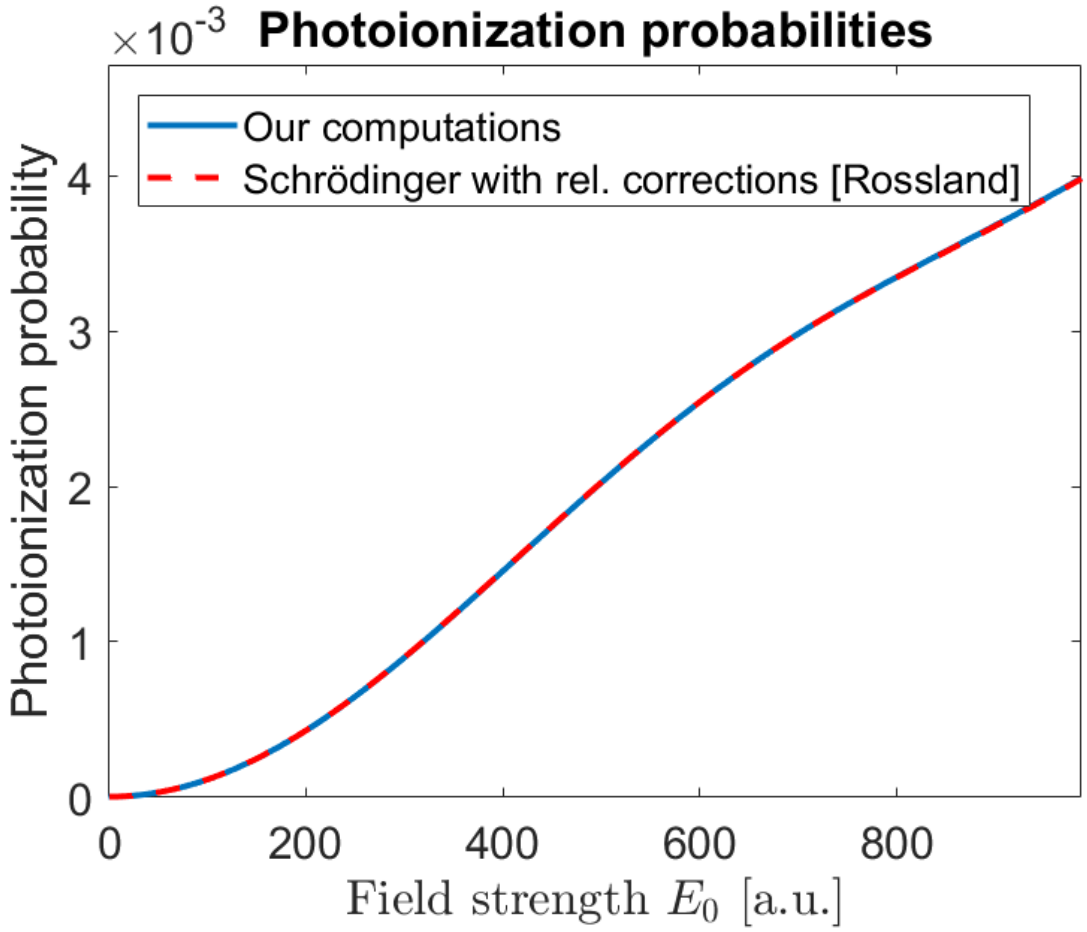


Figure 4.9: Comparison of probabilities for photoionization in the Dirac equation and in the Schrödinger equation with relativistic corrections[40] for different field strengths E_0.

Figure 4.9 shows great consistency between the photoionization probabilities for the Dirac equation and the Schrödinger equation with relativistic corrections. The consistency verifies the validity of our computations and that the difference in the photoionzation probabilities between the NR Schrödinger[40] and the Dirac equation observed in figure 4.8 indeed is a relativistic effect.

Chapter 5

Conclusions and Discussion

5.1 Summary and Conclusions

The relativistic Dirac equation for both with and without an external pulse has been solved numerically. The Dirac equation was firstly solved without an external pulse, and we avoided spurious states by using a dual kinetic balance basis. The solutions without an external pulse showed a decrease in energy for bound states compared to theoretical energies from the NR Schrödinger equation. This decrease was due to length contraction and it could be also observed for the wavefunctions ψ, which in the relativistic treatment showed greater probability to to be found near the nucleus.

The Dirac equation with an external pulse was solved by using Krylov subspace methods, which showed great convergence properties. We chose an electron in the ground state in a hydrogen atom as the initial state, and we investigated the probability of photoionization to positive pseudo continuum states. The photoionization probabilities were compared to corresponding probabilities yielded from the NR Schrödinger equation[40] and it showed, for the relativistic treatment, a shift towards higher energies in the differential probability energy distribution. The photonionzation probabilities' dependence on the electric field strength of the pulse was also investigated, and it showed that a relativistic treatment of the electron yielded a smaller photoionzation probability, especially for large field strengths.

The numerical results for bound states in the Dirac equation without an external pulse showed great consistency with theory. The Dirac equation with an external pulse is, on the hand, not analytically solvable, but the results showed good consistency with numerical results for both the Dirac equation in[41] and for the Schrödinger equation with relativistic corrections[40]. The validity of the numerical results for the Dirac equation with an external pulse is questionable

for high field strengths E_0, as the dipole approximation is not sufficient. It is suggested to go beyond the dipole approximation and investigate the implementation of the gauge transformation proposed in[22].

Bibliography

1. K. Sanderson. The most insense laser in the universe, nature, 2008. URL http://www.nature.com/news/2008/0802018/full/news.2008.608.html.

2. J. Amos. Xfel: Brilliant x-ray laser comes online, bbc, 2017.

3. A. Di Piazza, C. Müller, K. Z. Hatsagortsyan, and C. H. Keitel. Extremely high-intensity laser interactions with fundamental quantum systems. *Rev. Mod. Phys.*, 84:1177–1228, Aug 2012. doi: 10.1103/RevModPhys.84.1177. URL https://link.aps.org/doi/10.1103/RevModPhys.84.1177.

4. Ali Eftekhari. Boltzmann?s method of philosophy, January 2002. URL http://philsci-archive.pitt.edu/1719/.

5. M. Fowler. The photoelectric effect, university of virginia, 2008. URL http://galileo.phys.virginia.edu/classes/252/photoelectric_effect.html.

6. Cyberphysics. URL http://www.cyberphysics.co.uk/topics/atomic/Photoelectric_effect/Photoelectric_effect.html.

7. Albert Einstein. Concerning an heuristic point of view toward the emission and transformation of light. *American Journal of Physics*, 33(5):367, 1965.

8. Max Planck. On the law of distribution of energy in the normal spectrum. *Annalen der physik*, 4(553):1, 1901.

9. Niels Bohr. Xxxvii. on the constitution of atoms and molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 26(153):476–502, 1913.

10. M. Williams. What is bohr's atomic model?, universe today, 2016. URL http://www.universetoday.com/46886/bohrs-atomic-model/.

11. Peter Weinberger. Revisiting louis de broglie's famous 1924 paper in the philosophical magazine. *Philosophical magazine letters*, 86(7):405–410, 2006.

12. Erwin Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical review*, 28(6):1049, 1926.

13. Masanao Ozawa. Heisenberg's original derivation of the uncertainty principle and its universally valid reformulations. *arXiv preprint arXiv:1507.02010*, 2015.

14. James Clerk Maxwell. A treatise on electricity and magnetism, clarendon press, 1873.

15. Albert A Michelson and Edward W Morley. On the relative motion of the earth and of the luminiferous ether. *Sidereal Messenger, vol. 6, pp. 306-310*, 6:306–310, 1887.

16. Albert Einstein. On the electrodynamics of moving bodies. 1905.

17. Joseph Larmor. A dynamical theory of the electric and luminiferous medium. part iii. relations with material media. *Proceedings of the Royal Society of London*, 61(369-377):272–285, 1897.

18. James D. Bjorken and Sidney D. Drell. *Relativistic Quantum Mechanics*. McGraw-Hill Book Company, 1964.

19. Tommy Ohlsson. *Relativistic Quantum Physics*. Cambridge University Press, 2011.

20. Kenneth S Pitzer. Relativistic effects on chemical properties. *Accounts of chemical Research*, 12(8):271–276, 1979.

21. Tor Kjellsson, Sølve Selstø, and Eva Lindroth. Relativistic ionization dynamics for a hydrogen atom exposed to superintense xuv laser pulses. *Physical Review A*, 95(4):043403, 2017.

22. Tor Kjellsson, Morten Førre, Aleksander Skjerlie Simonsen, Sølve Selstø, and Eva Lindroth. Alternative gauge for the description of the light-matter interaction in a relativistic framework. *Phys. Rev. A*, 96:023426, Aug 2017. doi: 10.1103/PhysRevA.96.023426. URL https://link.aps.org/doi/10.1103/PhysRevA.96.023426.

23. Aleksander Skjerlie Simonsen and Morten Førre. Dipole-forbidden atomic transitions induced by superintense x-ray laser fields. *Phys. Rev. A*, 93:063425, Jun 2016. doi: 10.1103/PhysRevA.93.063425. URL https://link.aps.org/doi/10.1103/PhysRevA.93.063425.

24. S. Blanes, F. Casas, J.A. Oteo, and J. Ros. The magnus expansion and some of its applications. *Physics Reports*, 470(5):151 – 238, 2009. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2008.11.001. URL http://www.sciencedirect.com/science/article/pii/S0370157308004092.

25. Peter Arbeiz. Lecture notes: Chapter 10: Arnoldi and lanczsos algorithms, 2016. URL http://people.inf.ethz.ch/arbenz/ewp/Lnotes/chapter10.pdf.

26. Steven Duplij and Frans Klinkhamer. Levi-civita symbol. In *Concise Encyclopedia of Supersymmetry*, pages 227–227. Springer, 2004.

27. Hemmer. *Kvantemekanikk 5. utgave*. Fagbokforlaget, 2015.

28. Gordon Baym. *Lectures on quantum mechanics*. CRC Press, 2018.

29. V. Åsbrink A. Lawncizak. *On the 1932 Discovery of the Positron*. KTH Engineering Sciences, 2014.

30. G. Shaw F. Mandl. *Quantum Field Theory, second edition*. Wiley, 2010.

31. Dyson Freeman transcribed by David Derbes. *Advanced Quantum Mechanics, second edition*. World Scientific, 2011.

32. Jim Branson. Solution of the dirac equation for hydrogen, 2013. URL https://quantummechanics.ucsd.edu/ph130a/130_notes/node501.html.

33. Walter R Johnson. *Atomic structure theory*. Springer, 2007.

34. H. Bachau. *Applications of B-splines in atomic and molecular physics*. Rep. Prog. Phys 64, 1815, 2001.

35. Mathieu Lewin and Éric Séré. Spurious modes in dirac calculations and how to avoid them. In *Many-Electron Approaches in Physics, Chemistry and Mathematics*, pages 31–52. Springer, 2014.

36. V. M. Shabaev, I. I. Tupitsyn, V. A. Yerokhin, G. Plunien, and G. Soff. Dual kinetic balance approach to basis-set expansions for the dirac equation. *Phys. Rev. Lett.*, 93:130405, Sep 2004. doi: 10.1103/PhysRevLett.93.130405. URL https://link.aps.org/doi/10.1103/PhysRevLett.93.130405.

37. Yulian V. Vanne and Alejandro Saenz. Solution of the time-dependent dirac equation for multiphoton ionization of highly charged hydrogenlike ions. *Phys. Rev. A*, 85:033411, Mar 2012. doi: 10.1103/PhysRevA.85.033411. URL https://link.aps.org/doi/10.1103/PhysRevA.85.033411.

38. Vidal Alonso, Salvatore De Vincenzo, and Luigi Mondino. On the boundary conditions for the dirac equation. *European Journal of Physics*, 18(5):315, 1997.

39. Randolf Beerwerth and Heiko Bauke. Krylov subspace methods for the dirac equation. 188: 189, 03 2015.

40. Ingunn Koren Rossland. Interaction of the hydrogen atom with laser fields: A study of relativistic effects in ionization processes. MSc-thesis at University of Bergen, June 2018.

41. Andreas Skeidsvoll. Modelling interactions between hydrogen-like atoms and intense laser pulses with the dirac equation. MSc-thesis at University of Bergen, May 2018.

Appendix A

Appendix

A.1 Mathematical derivations in theory

This section contain proofs and derivations for various relations in the theory chapters.

A.1.1 $\left[H_D, \vec{L}\right]$

Proof of $[H_D, \vec{L}] = -c\vec{\alpha} \times \vec{\nabla}$:

$$\begin{aligned}\left[H_D, L_i\right] &= \left[-ic\vec{\alpha}\cdot\vec{\nabla} + c^2\beta - \frac{1}{|r|}, -i\left(\vec{r}\times\vec{\nabla}\right)_i\right] \\ &= -c\left[\vec{\alpha}\cdot\vec{\nabla}, \left(\vec{r}\times\vec{\nabla}\right)_i\right] - i\left[\frac{1}{|r|}, \left(\vec{r}\times\vec{\nabla}\right)_i\right]\end{aligned} \tag{A.1}$$

The second term will vanish:

$$\left[\frac{1}{|r|}, \left(\vec{r}\times\vec{\nabla}\right)_i\right] = \left[\frac{1}{|r|}\epsilon_{ijk} r_j \nabla_k - \epsilon_{ijk} r_j \nabla_k \frac{1}{|r|}\right] \tag{A.2}$$

As $|r| = \sqrt{r_1^2 + r_2^2 + r_3^2}$, equation (A.2) can be written as:

$$\begin{aligned}\left[\frac{1}{|r|}, \left(\vec{r}\times\vec{\nabla}\right)_i\right] &= \left[\frac{1}{|r|}\epsilon_{ijk} r_j \nabla_k - \frac{1}{|r|}\epsilon_{ijk} r_j \nabla_k + \frac{1}{2}\frac{1}{|r|^3}\epsilon_{ijk} r_j r_k\right] \\ &= \frac{1}{2}\frac{1}{|r|^3}\left(\vec{r}\times\vec{r}\right)_i \\ &= 0\end{aligned} \tag{A.3}$$

This result simplifies equation (A.1) to;

$$\begin{aligned}\left[H_D, L_i\right] &= -c\left[\vec{\alpha}\cdot\vec{\nabla}, \left(\vec{r}\times\vec{\nabla}\right)_i\right] \\ &= -c\left[\left(\vec{\alpha}\cdot\vec{\nabla}\right)\left(\vec{r}\times\vec{\nabla}\right)_i - \left(\vec{r}\times\vec{\nabla}\right)_i\left(\vec{\alpha}\cdot\vec{\nabla}\right)\right] \\ &= -c\left[\alpha_m\nabla_m\epsilon_{ijk} r_j \nabla_k - \epsilon_{ijk} r_j \nabla_k \alpha_m \nabla_m\right] \\ &= -c\left[\epsilon_{ijk}\alpha_j\nabla_k + \epsilon_{ijk} r_j \nabla_k \alpha_m \nabla_m - \epsilon_{ijk} r_j \nabla_k \alpha_m \nabla_m\right] \\ &= -c\left[\epsilon_{ijk}\alpha_j\nabla_k\right] \\ &= -c\left(\vec{\alpha}\times\vec{\nabla}\right)_i\end{aligned} \tag{A.4}$$

As L is a three component vector, $\vec{L} = \left(L_1, L_2, L_3\right)$, the commutator of H_D and $\vec{L}$ will be;

$$\left[H_D, \vec{L}\right] = -c\vec{\alpha}\times\vec{\nabla} \tag{A.5}$$

A.1.2 $\left[H_D, \vec{L}^2\right]$

Proof of $\left[H_D, \vec{L}^2\right] = 2i\Big((\vec{r}\cdot\alpha)(\vec{\nabla}\cdot\vec{\nabla}) - (\vec{r}\cdot\vec{\nabla})(\vec{\alpha}\cdot\vec{\alpha})\Big)$;

$$\begin{aligned}
\left[H_D, \vec{L}^2\right] &= \left[H_D, \vec{L}\cdot\vec{L}\right] \\
&= \left[H_D, L_i L_i\right] \\
&= H_D L_i L_i - L_i L_i H_D \\
&= \Big(L_i H_D + \left[H_D, L_i\right]\Big) L_i - L_i L_i H_D \\
&= L_i H_D L_i - L_i L_i H_D + \left[H_D, L_i\right] L_i \\
&= L_i \left[H_D, L_i\right] + \left[H_D, L_i\right] L_i \\
&\overset{(A.4)}{=} L_i\big(-\epsilon_{ijk}\alpha_j\nabla_k\big) + \big(-\epsilon_{ijk}\alpha_j\nabla_k\big) L_i \\
&= \big(-i\epsilon_{imn} r_m \nabla_n\big)\big(-\epsilon_{ijk}\alpha_j\nabla_k\big) + \big(-\epsilon_{ijk}\alpha_j\nabla_k\big)\big(-i\epsilon_{imn} r_m \nabla_n\big) \\
&= i\epsilon_{imn}\epsilon_{ijk}\big(r_m\nabla_n\alpha_j\nabla_k + \alpha_j\nabla_j r_m\nabla_n\big)
\end{aligned} \tag{A.6}$$

As $\epsilon_{imn}\epsilon_{ijk} = \delta_{m,j}\delta_{n,k} - \delta_{n,j}\delta_{m,k}$, equation (A.6) simplifies to;

$$\begin{aligned}
\left[H_D, \vec{L}^2\right] &= i\big(\delta_{m,j}\delta_{n,k} - \delta_{n,j}\delta_{m,k}\big)\big(r_m\nabla_n\alpha_j\nabla_k + \alpha_j\nabla_j r_m\nabla_n\big) \\
&= i\Big(r_j\nabla_n\alpha_j\nabla_n + \alpha_j\nabla_n r_j\nabla_n - r_m\nabla_j\alpha_j\nabla_m - \alpha_j\nabla_k r_k\nabla_j\Big) \\
&\overset{(\text{relabel})}{=} i\Big(r_j\nabla_n\alpha_j\nabla_n + \alpha_j\nabla_n r_j\nabla_n - r_j\nabla_n\alpha_n\nabla_j - \alpha_n\nabla_j r_j\nabla_n\Big) \\
&= i\Big(r_j\alpha_j\nabla_n\nabla_n + r_j\alpha_j\nabla_n\nabla_n + \alpha_n\nabla_n - r_j\nabla_n\alpha_n\nabla_j - r_j\alpha_n\nabla_j\nabla_n - \alpha_n\nabla_n\Big) \\
&= i\Big(2r_j\alpha_j\nabla_n\nabla_n - 2r_j\nabla_j\alpha_n\nabla_n\Big) \\
&= 2i\Big((\vec{r}\cdot\vec{\alpha})(\vec{\nabla}\cdot\vec{\nabla}) - (\vec{r}\cdot\vec{\nabla})(\vec{\alpha}\cdot\vec{\nabla})\Big)
\end{aligned} \tag{A.7}$$

A.1.3 $\left[H,\vec{S}\right]$

Proof of $[H_D,\vec{S}] = 2c\vec{\alpha}\times\vec{\nabla}$:

$$\left[H_D,\vec{S}_i\right] = \left[-ic\vec{\alpha}\cdot\vec{\nabla} + c^2\beta - \frac{1}{|r|},\frac{1}{2}\sigma_i\right] \tag{A.8}$$

Since $\left[\beta,\sigma_i\right] = 0 \quad$ for $i = 1,2,3$ in equation (A.8), equation (A.8) simplifies to;

$$\begin{aligned}\left[H_D,\vec{S}_i\right] &= -\frac{i}{2}c\left[\vec{\alpha}\cdot\vec{\nabla},\sigma_i\right] \\ &= -\frac{i}{2}c\left[\alpha_j\nabla_j\sigma_i - \sigma_i\alpha_j\nabla_j\right] \\ &= -\frac{i}{2}c\left[\nabla_j\left(\alpha_j\sigma_i - \sigma_i\alpha_j\right)\right] \\ &= -\frac{i}{2}c\nabla_j\left[\alpha_j,\sigma_i\right]\end{aligned} \tag{A.9}$$

It can be shown that $\left[\alpha_j,\sigma_i\right]$ is;

$$\left[\alpha_j,\sigma_i\right] = 2i\epsilon_{jik}\alpha_k = 2i\epsilon_{ikj}\alpha_k, \tag{A.10}$$

which simplifies equation (A.9) to;

$$\begin{aligned}\left[H_D,\vec{S}_i\right] &= c\nabla_j\epsilon_{ikj}\alpha_k \\ &= c\epsilon_{ikj}\alpha_k\nabla_j \\ &= c\left(\vec{\alpha}\times\vec{\nabla}\right)_i\end{aligned} \tag{A.11}$$

As S is a three component vector, $\vec{S} = \left(S_1,S_2,S_3\right)$, the commutator of H_D and $\vec{S}$ will be

$$\left[H_D,\vec{S}\right] = c\vec{\alpha}\times\vec{\nabla} \tag{A.12}$$

A.1.4 $\left[H_D, \vec{S}^2\right]$

Proof of $\left[H_D, \vec{S}^2\right] = 0$

$$\begin{aligned}
\left[H_D, \vec{S}^2\right] &= \left[H_D, \vec{S}\cdot\vec{S}\right] \\
&= \left[H_D, S_i S_i\right] \\
&= H_D S_i S_i - S_i S_i H_D \\
&= \Big(S_i H_D + \left[H_D, S_i\right]\Big) S_i - S_i S_i H_D \\
&= S_i\left[H_D, S_i\right] + \left[H_D, S_i\right] S_i \\
&\overset{(A.11)}{=} S_i \epsilon_{imn} c\alpha_m \nabla_n + \epsilon_{imn} c\alpha_m \nabla_n S_i \\
&= \frac{1}{2}\sigma_i \epsilon_{imn} c\alpha_m \nabla_n + \epsilon_{imn} c\alpha_m \nabla_n \frac{1}{2}\sigma_i \\
&= \frac{c}{2}\epsilon_{imn}\left[\sigma_i \alpha_m \nabla_n + \alpha_m \nabla_n \sigma_i\right] \\
&= \frac{c}{2}\epsilon_{imn}\left[\sigma_i \alpha_m \nabla_n + \alpha_m \sigma_i \nabla_n\right]
\end{aligned} \tag{A.13}$$

It can be shown that $\{\alpha_m, \sigma_i\} = \alpha_m\sigma_i + \sigma_i\alpha_m$ can be expressed simplified to;

$$\{\alpha_m, \sigma_i\} = \begin{bmatrix} 0_{2\times2} & 2\delta_{mi}\mathbb{1}_{2\times2} \\ 2\delta_{mi}\mathbb{1}_{2\times2} & 0_{2\times2} \end{bmatrix}, \tag{A.14}$$

which simplifies equation (A.13) to;

$$\begin{aligned}
\left[H_D, \vec{S}^2\right] &= \frac{c}{2}\epsilon_{imn}\Big[\sigma_i \alpha_m \nabla_n + \big(-\sigma_i\alpha_m + \{\alpha_m, \sigma_i\}\big)\nabla_n\Big] \\
&= \frac{c}{2}\epsilon_{imn}\{\alpha_m, \sigma_i\}\nabla_n \\
&= \frac{c}{2}\nabla_n \begin{bmatrix} 0_{2\times2} & 2\epsilon_{imn}\delta_{mi}\mathbb{1}_{2\times2} \\ 2\epsilon_{imn}\delta_{mi}\mathbb{1}_{2\times2} & 0_{2\times2} \end{bmatrix} \\
&= \frac{c}{2}\nabla_n \begin{bmatrix} 0_{2\times2} & 2\epsilon_{iin}\mathbb{1}_{2\times2} \\ 2\epsilon_{iin}\mathbb{1}_{2\times2} & 0_{2\times2} \end{bmatrix} \\
&= 0
\end{aligned} \tag{A.15}$$

A.1.5 Derivation of the Runge-Lenz vector K_R

Derivation of the Runge-Lenz vector K_R;
As one wants to differ between states that have parallel and anti-parallel spin $\vec{S}$ and angular momentum $\vec{L}$, it is natural to seek an operator which is of the form;

$$Q = \vec{\sigma} \cdot \vec{J} \tag{A.16}$$

Although it can be shown that $\Big[H_D, Q\Big]$ is non-zero. So let's try $Q = \beta\vec{\sigma} \cdot \vec{J}$ and investigate whether it is constant of motion or not.

$$\begin{aligned}
\Big[H_D, Q\Big] &= \Big[H_D, \beta\vec{\sigma} \cdot \vec{J}\Big] \\
&= H_D\beta\sigma_i J_i + \beta\sigma_i J_i H_D \\
&= H_D\beta\sigma_i J_i - \beta\sigma_i H_D J_i + \beta\sigma_i H_D J_i - \beta\sigma_i J_i H_D \\
&= \Big[H_D, \beta\sigma_i\Big] J_i + \beta\sigma_i \Big[H_D, J_i\Big]
\end{aligned} \tag{A.17}$$

$\Big[H_D, J_i\Big] = 0$ as shown in section 2.6, which simplifies equation (A.17) to;

$$\begin{aligned}
\Big[H_D, Q\Big] &= \Big[H_D, \beta\sigma_i\Big] J_i \\
&= \Big[-ic\vec{\alpha} \cdot \vec{\nabla} + c^2\beta - \frac{1}{|r|}, \beta\sigma_i\Big] J_i \\
&= \Big(-ic\Big[\vec{\alpha} \cdot \vec{\nabla}, \beta\sigma_i\Big] + c^2\Big[\beta, \beta\sigma_i\Big]\Big) J_i
\end{aligned} \tag{A.18}$$

The last term $[\beta, \beta\sigma_i]$ vanishes, as $\beta^2 = \mathbb{1}_4$ and $[\sigma_i, \beta] = 0$.
Equation (A.18) can then be written as;

$$\begin{aligned}
\Big[H_D, Q\Big] &= -ic\Big[\vec{\alpha} \cdot \vec{\nabla}, \beta\sigma_i\Big] J_i \\
&= -ic\Big[\alpha_j\nabla_j, \beta\sigma_i\Big] J_i \\
&= -ic\Big(\alpha_j\nabla_j\beta\sigma_j - \beta\sigma_i\alpha_j\nabla_j\Big) J_i \\
&= -ic\nabla_j\Big(\alpha_j\beta\sigma_i - \beta\sigma_i\alpha_j\Big) J_i
\end{aligned} \tag{A.19}$$

$\alpha_j\beta\sigma_i = -\beta\alpha_j\sigma_i$ as $\alpha_m\beta = -\beta\alpha_m$. Furthermore, $\alpha_j\sigma_i = -\sigma_i\alpha_j$ for $j \neq i$ by equation (A.14) and $\alpha_j\sigma_i = \sigma_i\alpha_j$ for $j = i$ by equation (A.10). These relations simplify equation (A.20) to;

$$\begin{aligned}\Big[H_D, Q\Big] &= -i\hbar c\nabla_i\big(-2\beta\sigma_i\alpha_i\big)J_i \\ &= -2ic\nabla_i\alpha_i\sigma_i\beta J_i \\ &= -2ic\nabla_i\alpha_i\sigma_i\beta\Big(\frac{1}{2}\sigma_i + (\vec{r}\times\vec{\nabla})_i\Big) \\ &= -ic\vec{\alpha}\cdot\vec{\nabla}\beta - 2ic\alpha_i\sigma_i\beta\vec{\nabla}\cdot(\vec{r}\times\vec{\nabla}) \\ &= -ic\vec{\alpha}\cdot\vec{\nabla}\beta,\end{aligned} \tag{A.20}$$

as the of the curl of a vector $\big(\vec{\nabla}\cdot(\vec{r}\times\vec{\nabla})\big)$ is zero. As $\Big[H_D, \frac{1}{2}\beta\Big] = -ic\vec{\alpha}\cdot\vec{\nabla}\beta$ too, the Runge-Lenz vector K_R will commute with the Dirac Hamiltonian H_D;

$$K_R = \beta\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\beta \tag{A.21}$$

A.1.6 $K_R^2 = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + 1\right)$

Derivation of $K_R^2\psi = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + 1\right)\psi$;
Starting by squaring the operator K_R given by equation (A.21)

$$\begin{aligned} K_R^2 &= \left(\beta\left(\vec{\sigma}\cdot\vec{J} - \frac{1}{2}\right)\right)^2 \\ &= \left(\beta\left(\vec{\sigma}\cdot\left(\frac{1}{2}\vec{\sigma} + \vec{L}\right) - \frac{1}{2}\right)\right)^2 \\ &= \left(\beta\left(\frac{\vec{\sigma}^2}{2} + \vec{\sigma}\cdot\vec{L} - \frac{1}{2}\right)\right)^2 \\ &= \left(\beta\left(\vec{\sigma}\cdot\vec{L} + 1\right)\right)^2 \\ &= \beta\left(\sigma_i L_i + 1\right)\beta\left(\sigma_j L_j + 1\right) \\ &= \sigma_i L_i \sigma_j L_j + \sigma_i L_i + \sigma_j L_j + 1 \\ &\overset{\text{(relabel)}}{=} \sigma_i \sigma_j L_i L_j + 2\sigma_i L_i + 1 \end{aligned} \tag{A.22}$$

Pauli matrices satisfy the following relation;

$$\sigma_i, \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k, \tag{A.23}$$

which transforms equation (A.22) to

$$\begin{aligned} K_R^2 &= L_i L_j\left(\delta_{ij} + i\epsilon_{ijk}\sigma_k\right) + 2\sigma_i L_i + 1 \\ &= \vec{L}^2 + i\vec{\sigma}\cdot\left(\vec{L}\times\vec{L}\right) + 2\sigma_i L_i + 1, \end{aligned} \tag{A.24}$$

where it can be shown that $\left(\vec{L}\times\vec{L}\right)_i = iL_i$, which transforms equation (A.24) to;

$$K_R^2 = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + 1\right) \tag{A.25}$$

A.1.7 $\vec{J}^2 = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + \frac{3}{4}\right)$

Derivation of $\vec{J}^2\psi = \left(\vec{L}^2 + \vec{\sigma}\cdot\vec{L} + \frac{3}{4}\right)\psi$;

We start with

$$\begin{aligned}\vec{J}^2 &= \left(\frac{1}{2}\sigma + \vec{L}\right)^2 \\ &= \left(\frac{1}{2}\sigma + \vec{L}\right)\cdot\left(\frac{1}{2}\sigma + \vec{L}\right) \\ &= \frac{\vec{\sigma}^2}{4} + \vec{\sigma}\cdot\vec{L} + \vec{L}^2 \\ &= \vec{L}^2 + \vec{\sigma}\cdot\vec{L} + \frac{3}{4}\end{aligned} \tag{A.26}$$

A.1.8 $K = \vec{\sigma} \cdot \vec{L} + 1$

Derivation of $K = \vec{\sigma} \cdot \vec{J} - \frac{1}{2} = \vec{\sigma} \cdot \vec{L} + 1$:

$$\begin{aligned} K &= \vec{\sigma} \cdot \vec{J} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \left(\vec{L} + \frac{1}{2}\sigma\right) - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + \frac{\vec{\sigma}^2}{2} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + \frac{3}{2} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + 1 \end{aligned} \tag{A.27}$$

A.1.9 $\vec{\sigma}\cdot\vec{p} = \frac{1}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\Big(-ir\frac{\partial}{\partial r} + i\vec{\sigma}\cdot\vec{L}\Big)$

$$\vec{\sigma}\cdot\vec{p} = \frac{1}{2}2\delta_{ij}\hat{x}_i\hat{x}_j\sigma_n p_n, \tag{A.28}$$

as $\delta_{ij}\hat{x}_i\hat{x}_j = \hat{x}_i\hat{x}_j = 1$. Furthermore,

$$\{\sigma_i,\sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}, \tag{A.29}$$

which transforms equation (A.28) to;

$$\begin{aligned}\vec{\sigma}\cdot\vec{p} &= \frac{1}{2}\big(\sigma_i\sigma_j + \sigma_j\sigma_i\big)\hat{x}_i\hat{x}_j\sigma_n\sigma_p \\ &\overset{\text{(relabel)}}{=} \sigma_i\sigma_j\hat{x}_i\hat{x}_j\sigma_n p_n\end{aligned} \tag{A.30}$$

As $\hat{x}_i = \frac{x_i}{r}$, equation(A.30) takes the form

$$\begin{aligned}\vec{\sigma}\cdot\vec{p} &= \frac{1}{r}\frac{\sigma_i x_i}{r}\big(\sigma_j\sigma_n x_j p_n\big) \\ &\overset{\text{(relabel)}}{=} \frac{1}{r}\frac{\sigma_i x_i}{r}\frac{1}{2}\big(\sigma_j\sigma_n x_j p_n + \sigma_n\sigma_j x_n p_j\big)\end{aligned} \tag{A.31}$$

Making use of the relation $\big[\sigma_i,\sigma_j\big] = 2i\epsilon_{ijk}\sigma_k \Rightarrow$

$$\begin{aligned}\vec{\sigma}\cdot\vec{p} &= \frac{1}{r}\frac{\sigma_i x_i}{r}\frac{1}{2}\Big(\sigma_j\sigma_n x_j p_n + \big(\sigma_j\sigma_n + 2i\epsilon_{njk}\sigma_k\big)x_n p_j\Big) \\ &= \frac{1}{r}\frac{\sigma_i x_i}{r}\Big(\frac{1}{2}\big(\sigma_j\sigma_n x_j p_n + \sigma_j\sigma_n x_n p_j\big) + i\epsilon_{njk}\sigma_j x_n p_j\Big) \\ &\overset{\text{(relabel)}}{=} \frac{1}{r}\frac{\sigma_i x_i}{r}\Big(\frac{1}{2}\big(\sigma_j\sigma_n x_j p_n + \sigma_n\sigma_j x_j p_n\big) + i\epsilon_{njk}\sigma_j x_n p_j\Big) \\ &= \frac{1}{r}\frac{\sigma_i x_i}{r}\Big(\frac{1}{2}\big(\sigma_j\sigma_n + \sigma_n\sigma_j\big)x_j p_n + e\sigma_k L_k\Big) \\ &\overset{\text{(A.14)}}{=} \frac{1}{r}\frac{\sigma_i x_i}{r}\Big(\frac{1}{2}2\delta_{jn}x_j p_n + i\sigma_k L_k\Big) \\ &= \frac{1}{r}\frac{\vec{\sigma}\cdot\vec{x}}{r}\Big(-ir\frac{\partial}{\partial r} + i\vec{\sigma}\cdot\vec{L}\Big)\end{aligned} \tag{A.32}$$

A.1.10 $K = \vec{\sigma} \cdot \vec{L} + 1$

Derivation of $K = \vec{\sigma} \cdot \vec{J} - \frac{1}{2} = \vec{\sigma} \cdot \vec{L} + 1$:

$$\begin{aligned} K &= \vec{\sigma} \cdot \vec{J} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \left(\vec{L} + \frac{1}{2}\sigma\right) - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + \frac{\vec{\sigma}^2}{2} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + \frac{3}{2} - \frac{1}{2} \\ &= \vec{\sigma} \cdot \vec{L} + 1 \end{aligned} \tag{A.33}$$