An advanced algorithm to solve the Dirac-Coulomb equation

Athanasios Gkrepis,[1, 2, a)] Odysseas Kosmas,[3] and Theocharis Kosmas[1]

[1]*Department of Physics, University of Ioannina, Ioannina, Greece*
[2]*Department of Informatics, University of Western Macedonia, Kastoria, Greece*
[3]*Conigital LTD, Birmingham-Coventry, UK*

[a)]Corresponding author: a.gkrepis@uowm.gr

Abstract. The Quantum Mechanical partial differential equations of Schrödinger, Klein-Gordon, Dirac, etc., are fundamental tools for describing the motion of particles as well as phenomena in the microcosm. However, usually they don't have analytical solutions and the use of advanced numerical methods are demanded for their solutions. In this work we derive an effective algorithm for solving the Dirac-Coulomb equation within the physics informed neural networks techniques. For a modern application, we test this code (in Python language) on the reproducibility of the Muonium, (μ^+e^-), antimuon-electron two-leptons system, which is of current importance in atomic, nuclear and particle physics.

INTRODUCTION

The structure of the Hydrogen atom as well as that of purely leptonic atoms, like the Muonium (μ^+e^-) we are interested in the present work, is well described within the framework of the quantum electrodynamics (QED). In the context of this theory, protons, electrons and muons are spin-1/2 Dirac particles (fermions) and their motion and dynamics could be described by the well known Dirac equation [1, 2]. This equation encompasses special relativity and quantum mechanical aspects. In the case of the two-body system Hydrogen atom, (p^+e^-), one way of the description is to consider the electron e^- moving inside the Coulomb field $V(r)$ created by the positively charged proton (p^+). In a similar way, for the Muonium we may assume that the μ^+ is producing a Coulomb field $V(r)$ inside of which the e^- is moving. Subsequently, the mathematical problem one is facing is to solve a system of two coupled ordinary differential Dirac equations of first order in the relative coordinate system of the two-particles.

Then, in both the above examples, the real problem one has to deal with is the solution of the radial Dirac equations for a single-particle (the e^-) moving in a central potential $V(r)$ produced by the positively charged lepton μ^+. This description is a good approximation provided that, if m_1 and m_2 are the masses of the proton (or anti-muon), and electron, respectively, then $m_1 \gg m_2$. Under such an assumption the mass parameter entering the Dirac equation is the m_e mass of the electron.

In this paper, we derive a numerical scheme for solving the Dirac-Coulomb radial equations by utilizing a physics informed neural network (PINN) [3]. The rest of the paper, is organized as follows. At first, we present the formalism of solving analytically the differential system of the radial coupled one-body Dirac equations. Next, we present the new numerical scheme and describe the steps of the physics informed neural networks employed as well as the application of the method in the Muonium purely leptonic atom. This method, is an extension of that derived previously [4] for solving the non-relativistic Schrödinger equation. Finally, we summarize our findings and discuss future perspectives of this project.

FORMULATION OF THE DIRAC-COULOMB EQUATION ANALYTIC SOLUTIONS

In atomic and molecular physics, for most calculations a common starting point is the "independent particle central field approximation". In the case of the muonic atoms and purely leptonic atom Mu (μ^+e^-), this approximation assumes that the electron moves independently within a 4-vector potential field A_μ of the form

$$A_0(r) = -e\phi(r), \quad r = |\mathbf{r}|, \quad \mathbf{A} = 0 \tag{1}$$

(equivalently $A_i = 0, i = 1,2,3$). Clearly $\phi(r)$ remains unchanged by any rotation about $r = 0$, but transforms as the component $A_0(x)$ of a 4-vector under Lorentz and Poincaré transformation such as boosts or translations. However, solutions in central potentials like that of Eq. (1) have a simple form which is convenient for further calculation. Under these assumptions on the 4-vector potential, the Dirac Hamiltonian H becomes

$$H = c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\cdot E_0 + V(r)I. \tag{2}$$

For any stationary (time-independent) solution with energy E the Dirac Hamiltonian satisfies the equation $H\Psi = E\Psi$.

For obtaining analytical solutions of the Dirac equation for the electron in central field, one can follow the procedure outlined by Bethe and Salpeter [1]. The Dirac equation for particle in the field $V(r)$ is the 4×4 system of equations,

$$H\Psi(\mathbf{r}) = [c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\cdot E_0 + V(r)I]\,\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{3}$$

where $E_0 = mc^2$ is the rest mass of the electron, $\boldsymbol{p} = -i\hbar\nabla = -i\hbar(\partial_x, \partial_y, \partial_z)$ denotes the momentum operator and $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, β denote the Dirac matrices representations

$$\boldsymbol{\alpha} = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix}, \qquad \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{4}$$

I the 2×2 identity matrix and σ the Pauli matrices. By setting

$$\Psi(\mathbf{r}) \equiv \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = \begin{bmatrix} f(r)\mathscr{Y}(\theta,\phi) \\ ig(r)\tilde{\mathscr{Y}}(\theta,\phi) \end{bmatrix} \tag{5}$$

where $\mathscr{Y}$ and $\tilde{\mathscr{Y}}$ denote the spin spherical harmonics, Eq. (3) can equivalently be written in matrix form as,

$$\begin{bmatrix} E_0 - E + V & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & -E_0 - E + V \end{bmatrix} \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{6}$$

Using the identity

$$\boldsymbol{\sigma}\cdot\mathbf{p} = \frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r^2}\left(-i\hbar r\frac{\partial}{\partial r} + i\boldsymbol{\sigma}\cdot\mathbf{L}\right) \tag{7}$$

after some elaboration, we obtain the equations

$$\begin{aligned} \boldsymbol{\sigma}\cdot\mathbf{p} f(r)\mathscr{Y}(\theta,\phi) &= i\hbar\frac{d}{dr}f(r)\tilde{\mathscr{Y}}(\theta,\phi) + i\frac{k+1}{r}\hbar f(r)\tilde{\mathscr{Y}}(\theta,\phi) \\ \boldsymbol{\sigma}\cdot\mathbf{p} g(r)\tilde{\mathscr{Y}}(\theta,\phi) &= i\hbar\frac{d}{dr}g(r)\mathscr{Y}(\theta,\phi) - i\frac{k-1}{r}\hbar g(r)\mathscr{Y}(\theta,\phi). \end{aligned} \tag{8}$$

Then, the known coupled Dirac equation providing the radial part of the wavefunction $\Psi(\mathbf{r})$ are

$$\begin{aligned} \frac{df}{dr} + \frac{1+\kappa}{r}f &= \frac{1}{\hbar c}(E_0 + E - V)g, \\ \frac{dg}{dr} + \frac{1-\kappa}{r}g &= \frac{1}{\hbar c}(E_0 - E + V)f, \end{aligned} \tag{9}$$

where f and g are the large (upper) and small (bottom) components respectively (some authors denote these functions the opposite way) and κ is the total spin quantum number. The system of Eqs. (9), can be simplified by using the transformation

$$\begin{aligned} f &= F/r, \\ g &= G/r, \end{aligned} \tag{10}$$

where $F(r)$ and $G(r)$ are the corresponding reduced radial wave functions, given by

$$\begin{aligned} \frac{dF}{dr} + \frac{\kappa}{r}F &= \frac{1}{\hbar c}(E_0 + E - V)G, \\ \frac{dG}{dr} - \frac{\kappa}{r}G &= \frac{1}{\hbar c}(E_0 - E + V)F. \end{aligned} \tag{11}$$

The latter differential system will be numerically solved in the next section. The system of Eqs. (9) admits solutions of the form [1, 2],

$$\begin{aligned} f &= \frac{(2\lambda)^{3/2}}{\Gamma(2\gamma+1)}\left[\frac{(1+\varepsilon)\Gamma(2\gamma+1+n')}{4N(N-\kappa)n'!}\right]^{1/2}(2\lambda r)^{\gamma-1}e^{-\lambda r}\left[(N-\kappa)\,{}_1F_1(-n', 2\gamma+1, 2\lambda r) - n'\,{}_1F_1(1-n', 2\gamma+1, 2\lambda r)\right] \\ g &= \frac{-(2\lambda)^{3/2}}{\Gamma(2\gamma+1)}\left[\frac{(1-\varepsilon)\Gamma(2\gamma+1+n')}{4N(N-\kappa)n'!}\right]^{1/2}(2\lambda r)^{\gamma-1}e^{-\lambda r}\left[(N-\kappa)\,{}_1F_1(-n', 2\gamma+1, 2\lambda r) + n'\,{}_1F_1(1-n', 2\gamma+1, 2\lambda r)\right] \end{aligned} \tag{12}$$

where

$$\varepsilon = \frac{E}{E_0}\,, \quad N = \frac{\alpha}{\sqrt{1-\varepsilon^2}}\,, \quad \gamma^2 = \kappa^2 - \alpha^2\,, \quad \lambda = \frac{E_0}{\hbar c}\sqrt{1-\varepsilon^2}\,, \quad n' = N - \gamma$$

and ${}_1F_1(A,C,z)$ denotes the confluent hypergeometric function of first kind.

Regarding the coefficients in front of f and g, they are chosen so as to satisfy the normalization condition,

$$\int_0^\infty \left[f^2(r) + g^2(r)\right] r^2\, dr = \int_0^\infty \left[F^2(r) + G^2(r)\right] dr = 1. \tag{13}$$

The energy eigenvalues which are determined simultaneously with the corresponding wave functions, are given by

$$E_{n,\kappa} = mc^2 \left[1 + \frac{\alpha^2}{\left(n - |\kappa| + \sqrt{\kappa^2 - \alpha^2}\right)^2}\right]^{-1/2}. \tag{14}$$

It is worth noting that compared to the Schrödinger radial equation's eigenvalues, the quantum number κ joins n (principal quantum number) in the calculation of E.

NUMERICAL SOLUTION UTILIZING PHYSICS INFORMED NEURAL NETWORK

Beginning with the radial Eqs. (11), if F and G are a pair of solutions and $r_i > 0$ is a finite sequence of positive numbers (i.e. a grid on the positive r-axis) then,

$$\sum_{i=0}^{s} \left\{ \left[\frac{dG(r_i)}{dr} + \frac{\kappa}{r_i} G(r_i) - \frac{1}{\hbar c}(E_0 + E - V(r_i))F(r_i)\right]^2 + \left[\frac{dF(r_i)}{dr} - \frac{\kappa}{r_i} F(r_i) - \frac{1}{\hbar c}(E_0 - E + V(r_i))G(r_i)\right]^2 \right\} = 0 \tag{15}$$

must also hold true.

Let us assume now that F and G are trial numerical solutions of Eqs. (9) given by,

$$\begin{aligned} F(r) &= c_f r e^{-\beta r} N(r, \vec{u}_f, \vec{v}_f, \vec{w}_f) \\ G(r) &= c_g\, r e^{-\beta r} N(r, \vec{u}_g, \vec{v}_g, \vec{w}_g) \end{aligned} \tag{16}$$

where $\beta > 0$ is an optimization constant related to $\lambda = m/\hbar c\sqrt{1-(E/mc^2)^2}$, $(c_f, c_g) \in \mathbb{R}^2$ and

$$N(r, \vec{u}, \vec{v}, \vec{w}) = \sum_{i=1}^{d} u_i\, a_f(v_i r + w_i) \tag{17}$$

is a one layer neural network with a_f denoting an activation function (e.g. sigmoid, tanh, ReLU).

Inserting these functions into the left hand side of Eq. (15), we define the error function, E_f for the grid r_i defined above as,

$$E_f(\vec{u}, \vec{v}, \vec{w}) = \sum_{i=0}^{s} \left\{ \left[\frac{dF(r_i)}{dr} + \frac{\kappa}{r_i} F(r_i) - \frac{1}{\hbar c}(E_0 + E + V)G(r_i)\right]^2 + \left[\frac{dG(r_i)}{dr} - \frac{\kappa}{r_i} G(r_i) - \frac{1}{\hbar c}(E_0 - E - V)F(r_i)\right]^2 \right\} \tag{18}$$

where $\vec{u} = (\vec{u_f}, \vec{u_g}), \vec{v} = (\vec{v_f}, \vec{v_g}), \vec{w} = (\vec{w_f}, \vec{w_g}) \in \mathbb{R}^{2d}$. One could divide the error function by $\int_0^\infty (F^2 + G^2)\, dr$ for avoiding the trivial zero solutions.

Now the task it to minimize the error function. For $\kappa = -1$, it has been solved using deterministic and stochastic methods for different systems [5, 6, 7].

As for the corresponding energy E, it can be calculated using the following formula,

$$E = \frac{\int_0^\infty V(r)[F^2(r) - G^2(r)]\, dr + E_0 \int_0^\infty \left(F^2 + G^2\right) dr}{\int_0^\infty [F^2(r) - G^2(r)]\, dr}. \tag{19}$$

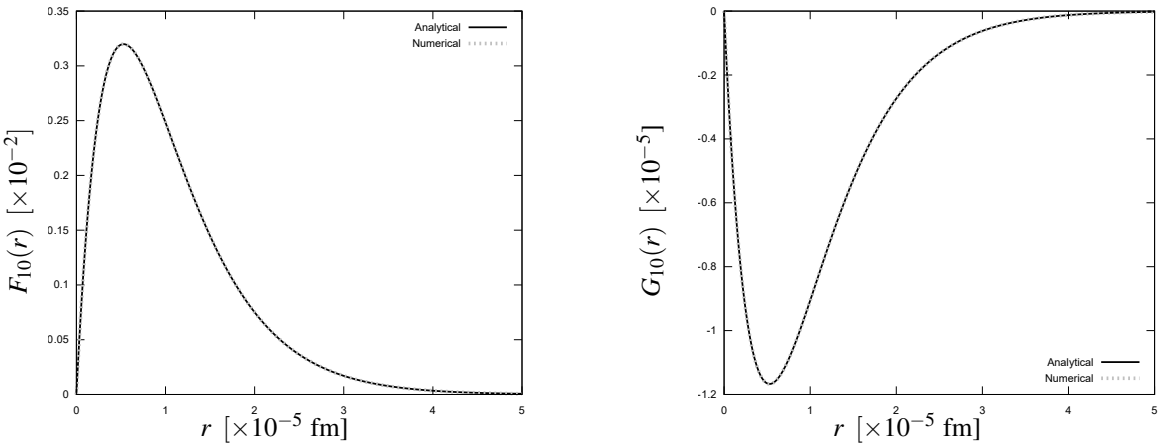

FIGURE 1. We see that, the analytical (straight line) and numerical (dashed line) solutions of F (left) and G (right) coincide.

In this work we utilized the L-BFGS-B method provided by the python scipy.optimize module [8], with 8 hidden units of tanh form, with $c_f = 1$ and obtained the results presented in the above figure for the ground state ($n = 1$, $l = 0$).

Furthermore, from the comparison the low-lying energy eigenvalues, corresponding to the wave functions obtained through the analytic and numerical methods for the Muonium exotic atom, we conclude that the performance of our algorithm is excellent.

SUMMARY, CONCLUSIONS AND OUTLOOK

We presented a numerical scheme relying on a Physics Informed Neural Network for solving the Dirac-Coulomb radial equations for a single electron system for which analytical solutions exist. Then, we applied this algorithm for obtaining the low-lying bound spectrum of the Muonium exotic atom.

Our next step is to solve numerically, the Dirac-Coulomb-Breit equation [9] which provides a fully relativistic description of two-body systems by incorporating, among other corrections, the two-particle Breit term.

ACKNOWLEDGMENTS

This work was financially supported from the Association for Advancement of Research on Open Problems (OPRA Association, Tel Aviv, Israel) in Nuclear Physics & Particle Physics, project No 83241/ELKE-UoI.

REFERENCES

1. H. A. Bethe and E. E. Salpeter, *Quantum mechanics of one- and two-electron atoms* (Dover Publications, Mineola, NY, 2014).
2. R. Boudet, *Relativistic Transitions in the Hydrogenic Atoms*, Springer Series on Atomic, Optical, and Plasma Physics, Vol. 52 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
3. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nat. Rev. Phys **3**, 422–440 (2021).
4. A. Gkrepis, O. Kosmas, D. Vlachos, and T. Kosmas, "Numerical solution of the schrödinger equation using neural networks in python," (to be published).
5. I. Lagaris, A. Likas, and D. Fotiadis, "Artificial neural network methods in quantum mechanics," Comput. Phys. Commun. **104**, 1–14 (1997).
6. T. S. Kosmas and I. E. Lagaris, "On the muon–nucleus integrals entering the neutrinoless $\mu^- \to e^-$ conversion rates," J. Phys. G: Nucl. Part. Phys. **28**, 2907–2920 (2002).
7. I. G. Tsoulos, O. Kosmas, and V. Stavrou, "Diracsolver: A tool for solving the Dirac equation," Comput. Phys. Commun. **236**, 237–243 (2019).
8. J. L. Morales and J. Nocedal, "Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization"," ACM Trans. Math. Softw. **38**, 1–4 (2011).
9. T. Papavasileiou, O. Kosmas, and T. Kosmas, "Mathematical formulation on the dirac-breit darwin equation for purely leptonic atoms," (to be published).