

An eight-component relativistic wave equation for spin- $\frac{1}{2}$ particles I

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1996 J. Phys. A: Math. Gen. 29 157

(<http://iopscience.iop.org/0305-4470/29/1/017>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.70

The article was downloaded on 02/06/2010 at 03:51

Please note that [terms and conditions apply](#).

An eight-component relativistic wave equation for spin- $\frac{1}{2}$ particles I

B A Robson and D S Staudte†

Department of Theoretical Physics, Research School of Physical Sciences and Engineering,
Australian National University, Canberra, ACT 0200 Australia

Received 2 May 1995

Abstract. An eight-component relativistic wave equation for spin- $\frac{1}{2}$ particles is derived. It is obtained from a spin- $\frac{1}{2}$ wave equation, which contains second-order derivatives in both space and time, by a procedure involving the linearization of the time derivative. It is in Hamiltonian form and is analogous to the two-component spin-0 equation, which features the use of an indefinite inner product in the description of the solution space. It is used to produce the relativistic bound-state energy eigenvalue spectrum and wavefunctions for the hydrogen atom.

1. Introduction

The discovery of electron spin in 1925 by Uhlenbeck and Goudsmit [1] led to attempts to describe it quantum mechanically. In 1927, Pauli [2] described the electron in terms of a two-component wavefunction satisfying a Schrödinger-type equation involving the 2×2 Pauli spin matrices. This theory is non-relativistic and, as Pauli emphasized, is therefore only approximate: one required an equation compatible with the special theory of relativity.

The first relativistic wave equation, written down by several people in 1926 [3], is the so-called Klein–Gordon equation (in the following we assume $c = 1 = \hbar$)

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + m^2 \right) \Psi = (\partial^\mu \partial_\mu + m^2) \Psi = 0. \quad (1)$$

The extension of equation (1) to the case of a classical external electromagnetic field is

$$(D^\mu D_\mu + m^2) \Psi = 0 \quad (2)$$

with D_μ the minimally coupled derivative [4].

This equation suffers from two basic flaws as far as the electron is concerned. First, it does not describe the spin of the electron, nor its magnetic moment. Secondly, it does not describe the bound-state energy levels of the hydrogen atom, as given by the Sommerfeld fine-structure formula, although it does give the Balmer formula. Also, at the time, there was no physical explanation for the non-positive definite nature of the density associated with the Klein–Gordon equation

$$\rho = j^0 = \frac{i}{m} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \frac{\partial \Psi^*}{\partial t} \Psi \right) - \frac{2eA_0}{m} \Psi^* \Psi \quad (3)$$

† Now at Fysik 3A, Chalmers Tekniska Högskola, 412 96 Göteborg, Sweden.

which thus cannot be interpreted as a probability density as in the non-relativistic case.

In 1928 Dirac [5] published his relativistic equation for the electron. He introduced 4×4 matrices γ^μ , where

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \quad (4)$$

to factorize the Klein–Gordon equation and obtained the Dirac equation

$$(i\gamma^\mu D_\mu - m)\Psi_D = 0 \quad (5)$$

linear in $\frac{\partial}{\partial t}$, which Dirac believed at the time was required to obtain a positive definite j^0 . The wavefunction Ψ_D has four components and the density $j^0 = \Psi_D^\dagger \Psi_D$ is positive definite and so can be interpreted as a probability density. In addition the Dirac equation describes particles with spin- $\frac{1}{2}$ and also the (almost) correct magnetic moment $e/(2m)$. The Sommerfeld fine-structure formula is also obtained for the bound-state energy levels of the hydrogen atom. Thus the Dirac equation was spectacularly successful in producing results in agreement with experiment at the time. However, there was one serious problem: the equation has negative energy solutions, which cannot easily be rejected. Indeed, Klein and Nishina [6] described Compton scattering in terms of the Dirac equation and their theory required the inclusion of the negative energy states in order to give the Thomson classical limit at low energies [7].

Initially it was suggested that the negative energy states were protons, but Dirac pointed out that protons had the wrong sign of charge. He then suggested that all the negative energy states are normally occupied and that ‘holes’ in this infinite sea of electrons act like a particle with positive energy and positive charge. Later he identified the holes with a new particle, which he termed the anti-electron. In 1932 Anderson [8] discovered the positron with a mass equal to that of the electron. This discovery was accepted by most people as a vindication of Dirac’s theory. Yet, the idea of a positron being a hole in an infinite sea of negative electrons was considered to be unsatisfactory to some, since it led to problems with infinities: the vacuum had a negative infinite zero point energy and an infinite zero point charge. Moreover, the scattering of a photon by an electron had become a many-body problem rather than a simple two-body problem.

The problem of the negative energy states was eventually overcome by the use of second quantization and quantum field theory, although this involved the use of a renormalization technique which is plagued by infinities. However, the resultant theory of quantum electrodynamics has been remarkably successful in making predictions in agreement with experiments, which later showed small discrepancies (e.g. the Lamb shift) with the Dirac theory.

However, the Dirac equation, as a relativistic wave equation, is often used in both atomic and nuclear physics, for cases where the external field is such that relativistic quantum mechanics is a useful approximation to quantum field theory. Here, the equation has the difficulty that certain predictions are not in agreement with nature. To be specific, the probability density is not a charge density and does not reflect the charge degree of freedom. Furthermore, the expectation value of the Hamiltonian in the case of no external field is not positive definite. In nature one sees the charge degree of freedom and only positive energies.

The purpose of the present paper is to show that there is another relativistic equation which also describes particles with spin- $\frac{1}{2}$. Moreover, this equation has an associated charge density which describes both positive and negative charged particles in a symmetrical way.

As well as the usual positive and negative energy electron solutions, the equation also has positive and negative energy positron solutions. Thus, the equation has solutions which match the physical description of positrons, i.e. having positive charge and positive energy. The equation is analogous to a two-component equation for spin-0 particles, obtained from the Klein–Gordon equation by linearizing the time derivative. The two-component spin-0 equation is discussed in section 2. The spin- $\frac{1}{2}$ analogue to it, which is the eight-component equation we wish to derive, is constructed in section 3. In section 4 the new equation is solved exactly to produce the bound-state energy eigenvalue spectra and wavefunctions for hydrogenic atoms. The method of solution is presented in detail so that it can be compared with the solution using the Dirac equation.

2. Hamiltonian form of the Klein–Gordon equation

The Klein–Gordon equation (1) contains second-order derivatives in both space and time. In order to obtain an equation in Hamiltonian form, it is necessary to linearize the time derivative. This can be done [9] by writing the Klein–Gordon wavefunction Ψ_{KG} in a two-component form:

$$\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Psi_{\text{KG}} + m^{-1}iD_0\Psi_{\text{KG}} \\ \Psi_{\text{KG}} - m^{-1}iD_0\Psi_{\text{KG}} \end{pmatrix} \quad (6)$$

so that

$$i\frac{\partial\Psi}{\partial t} = H\Psi \quad (7)$$

with

$$H = -\frac{1}{2m}D^2 \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} + m \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + eA_0\mathbf{1}_2 \quad (8)$$

where D and D_0 are the minimally coupled three-dimensional space and time derivatives respectively, and $\mathbf{1}_2$ is the 2×2 unit matrix. We call equation (7) the two-component spin-0 equation [9] and denote the corresponding wavefunction by Ψ . Feshbach and Villars [9] show that the writing of the Klein–Gordon equation in this (Hamiltonian) form opens the possibility of developing a quantum mechanical formalism for the equation, providing an indefinite inner product is used. The density (3) becomes

$$j^0 = \phi^*\phi - \chi^*\chi = \Psi^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Psi \quad (9)$$

which is not positive definite but can readily be interpreted as a charge density. Moreover, the indefinite inner product provides an interpretation in which only positive physical energies are obtained.

3. Derivation of the eight-component equation

In section 2 the two-component spin-0 equation is given by equation (7) with the Hamiltonian defined by equation (8). It is obtained from the Klein–Gordon equation by linearizing the time derivative.

An analogous equation for spin- $\frac{1}{2}$ particles may be constructed by linearizing the time derivative of a corresponding second-order equation which includes the appropriate spin terms. Such a second-order equation, which is manifestly covariant, is

$$((\gamma^\mu D_\mu)^2 + m^2 \mathbf{1}_4) \Psi = 0 \quad (10)$$

where the γ^μ are the usual Dirac matrices, and the wavefunction Ψ has four components. Equation (10), also known as the Feynman–Gell-Mann equation [10], may be written in the form

$$\left((D^\mu D_\mu + m^2) \mathbf{1}_4 + \frac{e}{2} \sigma^{\mu\nu} F_{\mu\nu} \right) \Psi = 0 \quad (11)$$

where $\sigma^{\mu\nu} = (i/2)[\gamma^\mu, \gamma^\nu]$ and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The final term contains the spin interaction in the presence of a classical external electromagnetic field. This term is absent for spin-0 particles. For free particles, equation (11) becomes the free particle Klein–Gordon equation for each component of the wavefunction.

Using the Weyl representation of the gamma matrices:

$$\gamma^0 = \begin{pmatrix} \mathbf{0} & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} \end{pmatrix} \quad \gamma = \begin{pmatrix} \mathbf{0} & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix} \quad (12)$$

equation (10) separates into two equations involving two-component wavefunctions Ψ_\pm :

$$(D_0^2 \mathbf{1}_2 - (\boldsymbol{\sigma} \cdot \mathbf{D})^2 \pm [(\boldsymbol{\sigma} \cdot \mathbf{D}), D_0 \mathbf{1}_2] + m^2 \mathbf{1}_2) \Psi_\pm = 0. \quad (13)$$

Equations (13) can be written

$$(D_0^2 \mathbf{1}_2 - D^2 \mathbf{1}_2 \pm i e \boldsymbol{\sigma} \cdot (\mathbf{E} \pm i \mathbf{B}) + m^2 \mathbf{1}_2) \Psi_\pm = 0 \quad (14)$$

where \mathbf{E} and \mathbf{B} are the usual electromagnetic field intensities. Linearizing equations (14) in the time derivative, analogously to equation (6), gives two four-component equations in Hamiltonian form with Hamiltonians

$$H_\xi = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \otimes \frac{1}{2m} (-D^2 \mathbf{1}_2 + i e \boldsymbol{\sigma} \cdot (\mathbf{E} + i \mathbf{B})) + m \begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} + e A_0 \mathbf{1}_4 \quad (15)$$

$$H_\eta = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \otimes \frac{1}{2m} (-D^2 \mathbf{1}_2 - i e \boldsymbol{\sigma} \cdot (\mathbf{E} - i \mathbf{B})) + m \begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} + e A_0 \mathbf{1}_4 \quad (16)$$

respectively, with \otimes the usual Kronecker (direct) product. The subscripts ξ and η are used here in place of $+$ and $-$ respectively. A more formal justification of the above derivation of equations (15) and (16) in terms of irreducible spinor representations ε^α and η_β of the proper Lorentz group is presented in the following paper [11]. The Hamiltonians H_ξ and H_η contain the pseudoscalar term $\boldsymbol{\sigma} \cdot \mathbf{E}$ and interchange under spatial inversion. In order to have an equation invariant under spatial inversion, it is necessary to consider an eight-component wavefunction $\Psi = (\Psi_\varepsilon, \Psi_\eta)^T$, where Ψ_ε and Ψ_η are four-component wavefunctions which transform into each other under spatial inversion. Thus one obtains the following equation for the eight-component wavefunction Ψ :

$$\begin{pmatrix} (i \frac{\partial}{\partial t} \mathbf{1}_4 - H_\xi) & \mathbf{0} \\ \mathbf{0} & (i \frac{\partial}{\partial t} \mathbf{1}_4 - H_\eta) \end{pmatrix} \begin{pmatrix} \Psi_\varepsilon \\ \Psi_\eta \end{pmatrix} = 0. \quad (17)$$

This eight-component equation is the spin- $\frac{1}{2}$ analogue to the two-component spin-0 equation. The conserved current and (indefinite) inner product for equation (17) are discussed in the following paper [11].

4. The hydrogen atom

Historically, the success of the Dirac equation in explaining the spectrum of the hydrogen atom was important for its acceptance as a relativistic wave equation. Nowadays the Dirac equation is used widely as a starting equation for more complex atomic calculations. If the eight-component equation derived in the previous section is a valid relativistic wave equation, it should also produce the bound-state energy eigenvalue spectrum for the hydrogen atom. Moreover, it is interesting to compare the solution method, spectrum and wavefunctions with that of the Dirac equation.

The general procedure used to obtain the hydrogen atom spectrum and wavefunctions is the following [12, 13]. One starts with the Dirac equation and adds a minimally coupled Coulomb potential, $A_0 = -Ze/r$, $\mathbf{A} = \mathbf{0}$. The Z in the Coulomb potential refers to a nucleus of Z protons, so that this solution also applies to hydrogen-like ions. Spherical polar coordinates are used, and it is possible to separate the radial and angular equations. The angular dependence is given by linear combinations of spherical harmonics coupled to two-component spinors while the radial functions are linear combinations of confluent hypergeometric functions multiplied by a decaying exponential function of the radial coordinate. The confluent hypergeometric function itself is defined in terms of a series and satisfies a second-order differential equation. Since the radial equations of the Dirac are coupled first-order equations, rather than a single second-order equation, it is not possible to write down the solution as a confluent hypergeometric function just by inspection. Instead a power series solution ansatz is tried and the series is found to match that of a certain linear combination of confluent hypergeometric functions.

A similar procedure will be applied to obtain the exact relativistic solution to the hydrogen atom for the eight-component equation. Not only will the hydrogen atom be solved, but the solution for spinless atoms will also be obtained, because the Hamiltonians for spin- $\frac{1}{2}$ (equations (15) and (16)) and spin-0 (equation (8)) are so similar. Obviously this is not possible using the Dirac equation. Given that the effects of the spin are only a relatively small correction to the spectrum, it would seem natural to have a method of solution where the effects of spin appear in the equation as a small extra term.

It turns out that the energy spectrum can be obtained with the use of only one of the four-component Hamiltonians (15) or (16). The Hamiltonian (15) is used with a Coulomb potential. One of equations (14) has been solved already to produce the hydrogen atom spectrum [14], but it is necessary to use the Hamiltonian (15) (or (16)), to see how the method of solution compares with the Dirac case. It is convenient to start with (15) with \hbar and c put in to facilitate the transition to atomic units. The atomic units are those of Bethe and Salpeter [12]. The equation $(i\frac{\partial}{\partial t}\mathbf{1}_4 - H_\xi)\Psi_\xi = 0$ with the Hamiltonian H_ξ given by (15) becomes, written as two two-component equations in atomic units with $\alpha = e^2/(\hbar c)$,

$$\left(\left(-\nabla^2 + \frac{2}{\alpha^2} - 2E - \frac{2Z}{r} \right) \mathbf{1}_2 - i\alpha Z \frac{\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}}{r^2} \right) \phi + \left(-\nabla^2 \mathbf{1}_2 - i\alpha Z \frac{\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}}{r^2} \right) \chi = 0 \quad (18)$$

$$\left(\nabla^2 \mathbf{1}_2 + i\alpha Z \frac{\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}}{r^2} \right) \phi + \left(\left(\nabla^2 - \frac{2}{\alpha^2} - 2E - \frac{2Z}{r} \right) \mathbf{1}_2 + i\alpha Z \frac{\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}}{r^2} \right) \chi = 0. \quad (19)$$

The Hamiltonian (16) gives the same equations except that the sign of the $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ term is changed. The spin-0 Hamiltonian (8) gives similar equations except that the $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ term is missing, and the two equations are each for only one component.

∇^2 is written in spherical polar coordinates and then the radial and angular variables are separated by using wavefunctions which are eigenfunctions of the angular operators in the

equation. The angular part of the ∇^2 operator has as eigenfunctions the spherical harmonics $Y_{lm}(\theta, \phi)$. For spin- $\frac{1}{2}$ particles the spherical harmonics must be coupled to two-component spinors in the standard way to give the spherical spinors $Y_{lj}^m(\theta, \phi)$ [13, 15]. The $\sigma \cdot \hat{r}$ operator interchanges the spherical spinors with $l = j \pm \frac{1}{2}$ and $l' = j \mp \frac{1}{2}$ (page 928 of [13]). Let Ω be the angular operator $L^2 \mathbf{1}_2 - i\alpha Z \sigma \cdot \hat{r}$. The eigenfunctions of Ω are

$$\Theta_{l,l',j,m}(\theta, \phi) = Y_{lj}^m + i\kappa Y_{l'j}^m \quad (20)$$

with

$$Z\alpha\kappa = \pm \left((j + \frac{1}{2}) - ((j + \frac{1}{2})^2 - (Z\alpha)^2)^{1/2} \right) \quad l = j \pm \frac{1}{2}. \quad (21)$$

The eigenvalues of Ω are

$$\lambda = (j + \frac{1}{2})^2 \pm ((j + \frac{1}{2})^2 - (Z\alpha)^2)^{1/2} \quad l = j \pm \frac{1}{2}. \quad (22)$$

For small Z , λ reduces to $l(l+1) + O(Z\alpha)$ as the spin correction is now small, and $\kappa \approx (Z\alpha)/(2j+1)$ so that the angular eigenfunctions are approximately Y_{lj}^m which are the non-relativistic (Pauli) eigenfunctions with orbital angular momentum l .

Writing the wavefunction as

$$\begin{pmatrix} \phi(r, \theta, \phi) \\ \chi(r, \theta, \phi) \end{pmatrix} = \frac{1}{r} \begin{pmatrix} f(r) \\ g(r) \end{pmatrix} \otimes \Theta_{(l,l',j,m)}(\theta, \phi) \quad (23)$$

gives the radial equations

$$\left(-\frac{\partial^2}{\partial r^2} + \frac{\lambda}{r^2} + \frac{2}{\alpha^2} - 2E - \frac{2Z}{r} \right) f(r) + \left(-\frac{\partial^2}{\partial r^2} + \frac{\lambda}{r^2} \right) g(r) = 0 \quad (24)$$

$$\left(\frac{\partial^2}{\partial r^2} - \frac{\lambda}{r^2} \right) f(r) + \left(\frac{\partial^2}{\partial r^2} - \frac{\lambda}{r^2} - 2E - \frac{2}{\alpha^2} - \frac{2Z}{r} \right) g(r) = 0. \quad (25)$$

In the spin-zero case one obtains the same radial equations except that $\lambda = l(l+1)$. Hence the radial equations will be solved simultaneously for spin- $\frac{1}{2}$ and spin-0.

Consider the $r \rightarrow \infty$ limit. The potential terms are then negligible and the radial equations become

$$\left(\frac{\partial^2}{\partial r^2} + 2E - \frac{2}{\alpha^2} \right) f + \frac{\partial^2}{\partial r^2} g = 0 \quad (26)$$

$$\frac{\partial^2}{\partial r^2} f + \left(\frac{\partial^2}{\partial r^2} - 2E - \frac{2}{\alpha^2} \right) g = 0. \quad (27)$$

This gives the asymptotic form of both f and g as

$$ae^{-|\Lambda|r} \quad |\Lambda| = \frac{1}{\alpha} \sqrt{1 - (E\alpha^2)^2}. \quad (28)$$

Here, a is a constant, different for f and g . This is the same asymptotic form as for the Dirac equation. The factor $e^{-|\Lambda|r}$ is eliminated from equations (24) and (25) by writing

$$f(r) = e^{-|\Lambda|r} y(r) \quad g(r) = e^{-|\Lambda|r} z(r). \quad (29)$$

Changing variables to $\rho = 2|\lambda|r$ and taking sums and differences of (24) and (25) gives

$$4|\Lambda|^2 \left(\frac{\partial^2}{\partial \rho^2} - \frac{\partial}{\partial \rho} + \frac{1}{4} - \frac{\lambda}{\rho^2} - \frac{1}{\alpha^2 4|\Lambda|^2} \right) (y(\rho) + z(\rho)) + \left(E + \frac{2|\Lambda|Z}{\rho} \right) (y(\rho) - z(\rho)) = 0 \quad (30)$$

$$\left(E + \frac{2|\Lambda|Z}{\rho} \right) (y(\rho) + z(\rho)) - \frac{1}{\alpha^2} (y(\rho) - z(\rho)) = 0. \quad (31)$$

To solve (30) and (31) a power series solution is attempted:

$$y(\rho) = \rho^\gamma \sum_{n=0}^{\infty} a_n \rho^n \quad (32)$$

$$z(\rho) = \rho^\gamma \sum_{n=0}^{\infty} b_n \rho^n. \quad (33)$$

If the power series extends to ∞ then the wavefunctions will not be normalizable. Both here and in the Dirac case one postulates that the series must therefore terminate at a certain value of n , and the condition that it terminates gives the energy spectrum. By substituting the power series solution ansatz into (30) and (31) and equating the coefficients of the lowest powers of ρ , γ is obtained. It is found that

$$\gamma(\gamma + 1) - \lambda + (Z\alpha)^2 = 0 \quad (34)$$

which gives

$$\gamma = -\frac{1}{2} + \left(\left(j + \frac{1}{2} \right)^2 - (Z\alpha)^2 \right)^{1/2} \pm \frac{1}{2} \quad l = j \pm \frac{1}{2}. \quad (35)$$

For spin-0

$$\gamma = -\frac{1}{2} + \left(\left(l + \frac{1}{2} \right)^2 - (Z\alpha)^2 \right)^{1/2}. \quad (36)$$

Equation (34) gives two values of γ for each value of λ , but, as in the Dirac case, only the one with the positive sign before the square root is chosen, otherwise for high j or l the wavefunctions will diverge strongly at the origin. General relations between the coefficients of the power series solution are, defining $c_n = a_n + b_n$ and $d_n = a_n - b_n$,

$$Ec_n + 2|\Lambda|Zc_{n+1} - \frac{1}{\alpha^2}d_n = 0 \quad (37)$$

$$((\gamma + n)(\gamma + n + 1) - \lambda)c_{n+1} - (\gamma + n)c_n + \frac{Z}{2|\Lambda|}d_n - \frac{E^2\alpha^2}{4|\Lambda|^2}c_{n-1} + \frac{E}{4|\Lambda|^2}d_{n-1} = 0. \quad (38)$$

From these equations it is simple to obtain both the energy spectrum and the wavefunctions. Rewrite (37) and (38) as

$$2|\Lambda|Zc_n = - \left(Ec_{n-1} - \frac{1}{\alpha^2}d_{n-1} \right) \quad (39)$$

$$((\gamma + n)(\gamma + n + 1) - \lambda)c_{n+1} - (\gamma + n)c_n + \frac{Z}{2|\Lambda|}d_n - \frac{E\alpha^2}{4|\Lambda|^2} \left(Ec_{n-1} - \frac{1}{\alpha^2}d_{n-1} \right) = 0. \quad (40)$$

Equation (39) is inserted into (40) to eliminate the $n - 1$ terms, and then (37) is used to eliminate the d_n term, leaving an equation for c_{n+1} and c_n only. This is

$$((\gamma + n)(\gamma + n + 1) - \lambda + (Z\alpha)^2)c_{n+1} + \left(\frac{ZE\alpha^2}{|\Lambda|} - (\gamma + n) \right) c_n = 0. \quad (41)$$

The first term can be simplified by using (34)

$$n(n + 2\gamma + 1)c_{n+1} + \left(\frac{ZE\alpha^2}{|\Lambda|} - n \right) c_n = 0. \quad (42)$$

It was mentioned before that there will be an $n = n' \geq 1$ where the series will terminate, and thus we set n so that the $n' + 1$ and higher coefficients are equal to zero. Hence (42) gives

$$n' = \frac{ZE\alpha^2}{|\Lambda|} - \gamma. \quad (43)$$

With the value for $|\Lambda|$ inserted one obtains a condition for E in terms of n' , which is

$$E = \frac{1}{\alpha^2} \left(1 + \frac{(Z\alpha)^2}{(n' + \gamma)^2} \right)^{-1/2}. \quad (44)$$

This is the same energy spectrum that the Dirac equation gives. The n' here begins at 1, whereas the Dirac n' begins at 0. However, unlike the γ for the Dirac equation, (35) has two solutions $\gamma = \gamma_{(\text{Dirac})} - 1$, $\gamma = \gamma_{(\text{Dirac})}$. The lowest energy state is for $n' = 1$ and here one will have $\gamma = \gamma_{(\text{Dirac})} - 1$ because it will be the $l = 0$ state and $l = j - \frac{1}{2}$, which picks the sign of λ in (22), and hence the lower value of γ . So actually the spectrum begins at the same value and then increases in correspondence with the Dirac spectrum. In both spectra there is a degeneracy for different values of n' and γ , as long as $n' + \gamma$ remains constant. One can show that the values of the quantum numbers j, l, m for each energy state are the same in both spectra, as one would expect given that they are physically verified. The only difference between the two equations is in the wavefunctions, which include polynomials of the form $\rho^\gamma \sum_{n=0}^{n'} a_n \rho^n$. It is not surprising to expect the wavefunctions to deviate somewhat as the Hamiltonian is different. To obtain the wavefunctions, consider (42). A confluent hypergeometric function $F(a, b; x)$ is defined as

$$F(a, b; x) = 1 + \frac{ax}{b} + \frac{a(a+1)x^2}{b(b+1)2!} + \dots \quad (45)$$

The coefficient f_q of x^q is related to the coefficient of x^{q+1} by

$$f_{q+1} = \frac{(a+q)}{(b+q)(q+1)} f_q. \quad (46)$$

Equation (42) is easily rewritten as

$$c_{n+1} = \frac{(-(ZE\alpha^2/|\Lambda|) - \gamma) + n}{(2\gamma + 1 + n)(n)} c_n = \frac{(-n' + n)}{(2\gamma + 1 + n)(n)} c_n. \quad (47)$$

This is not quite in the form (46) due to (n) and not $(n + 1)$ appearing in the denominator. However if a factor of ρ is taken from the power series $\sum_{n=0}^{\infty} c_n \rho^n$ then one indeed obtains a confluent hypergeometric function. The origin of this is that although $a_0 \neq 0$ and $b_0 \neq 0$, it is easily shown that $c_0 = 0$. Hence $c_1 \rho$ is the first non-zero term in the power series $y + z$ and becomes the first term in the corresponding confluent hypergeometric function. It can also be shown that for $n \leq n'$, $c_n \neq 0$, ($n \geq 1$), so that there is indeed a non-trivial power series. n' must be at least 1: if it were zero then it can be shown that $a_n = 0 = b_n$ and thus there is only the trivial solution.

Taking out a factor of ρ gives

$$c_n = \frac{(-n' + n)}{(2\gamma + 1 + n)(n)} c_{n-1} \tag{48}$$

which leads to

$$(y(\rho) + z(\rho)) = C \rho^{\gamma+1} F(1 - n', 2\gamma + 2; \rho) \tag{49}$$

where C is a (normalization) constant. Now consider (39). The n th term of the power series

$$E(y(\rho) + z(\rho)) - \frac{1}{\alpha^2}(y(\rho) - z(\rho))$$

equals the $(n + 1)$ th term of the power series $-2|\Lambda|Z(y(\rho) + z(\rho))$ and since the first term of the power series $(y(\rho) + z(\rho))$ is zero and hence has been removed, then

$$E(y(\rho) + z(\rho)) - \frac{1}{\alpha^2}(y(\rho) - z(\rho)) = -2|\Lambda|ZC\rho^\gamma F(1 - n', 2\gamma + 2; \rho). \tag{50}$$

From equations (49) and (50) $y(\rho)$ and $z(\rho)$ are obtained

$$\begin{Bmatrix} y(\rho) \\ z(\rho) \end{Bmatrix} = \frac{C}{2} \rho^\gamma F(1 - n', 2\gamma + 2; \rho) \begin{Bmatrix} (1 \pm \varepsilon)\rho \pm 2(Z\alpha)\sqrt{1 - \varepsilon^2} \end{Bmatrix} \tag{51}$$

with $\varepsilon = E\alpha^2$. In (51) the coefficients of the lowest power of ρ in the polynomials are of opposite sign for y and z . This explains why the result $a_0 + b_0 = c_0 = 0$, $a_0 \neq 0$, $b_0 \neq 0$ was derived earlier. The general form of the wavefunction $\Psi_\xi = \Psi_{jm}(r, \theta, \phi)$ is

$$\begin{aligned} \Psi_{jm}(r, \theta, \phi) &= CC' |\Lambda| e^{-|\Lambda|r} (2|\Lambda|r)^\gamma r^{-1} F(1 - n', 2\gamma + 2; 2|\Lambda|r) \\ &\times \begin{pmatrix} (1 + \varepsilon)r + \alpha^2 Z \\ (1 - \varepsilon)r - \alpha^2 Z \end{pmatrix} \otimes \Theta_{l,l',j,m}(\theta, \phi) \end{aligned} \tag{52}$$

with C and C' normalization constants for the radial and angular integrals respectively. These will be determined in the following paper [11], as the inner product needs to be derived before the normalization can be done. There it will be seen that the inner product involves both Ψ_ξ and $\Psi_{\tilde{\eta}}$, showing that the full eight-component equation is necessary in the analysis of the solutions. If $H_{\tilde{\eta}}$ is used for the hydrogen atom instead of H_ξ , then the same energy spectrum is obtained. The radial wavefunctions are also the same and the angular functions differ only in that the overall sign of κ (equation (21)) changes. This is what is expected given that $H_{\tilde{\eta}}$ is obtained from H_ξ by spatial inversion. The wavefunctions derived here are compared with the Dirac wavefunctions (page 69 of [12]) in [11]. There are differences in the radial wavefunctions, in particular the confluent hypergeometric functions. It will be interesting to see if any physical predictions deviate as a result of this.

To obtain the spin-0 results, all that is necessary is to replace γ from equation (35) by (36), and the radial functions will be analogous to (51). The angular functions are obtained by replacing $\Theta_{l,l',j,m}(\theta, \phi)$ by $Y_{lm}(\theta, \phi)$. Finally the energy spectrum is the same as (44) with γ from equation (36).

5. Concluding remarks

The eight-component equation, which is the spin- $\frac{1}{2}$ analogue to the two-component spin-0 equation, has been constructed. The derivation is via a generalization of the Klein–Gordon equation to include spin. It has been shown that the equation gives the same bound-state energy eigenvalue spectra for hydrogenic atoms as the Dirac equation, although the wavefunctions are slightly different, corresponding to a different Hamiltonian.

Acknowledgments

The authors wish to thank J M C F Govaerts for numerous invaluable discussions and insights into this work. They also wish to thank F Bastianelli, D J Daniel, R Dreizler, K Kumar, B Laurent and B H J McKellar for useful comments. One of the authors (DSS) is grateful to C Jarlskog and R Lindkvist for arranging a fruitful six month visit during 1992 to Fysikum, Stockholms Universitet, Sweden. He is also indebted to the Australian Government for an Australian Postgraduate Research Award and to A M Mårtensson Pendrill and the Swedish Institute for financial support.

References

- [1] Uhlenbeck G E and Goudsmit S 1925 *Naturwiss.* **13** 953
- [2] Pauli W 1927 *Z. Phys.* **43** 601
- [3] Klein O 1926 *Z. Phys.* **37** 895
Schrödinger E 1926 *Ann. Phys.* **38** 109
Fock V 1926 *Z. Phys.* **38** 242
de Donder T and van den Dungen H 1926 *C. R. Acad. Sci. Sér. Sci. Math. Astron. Phys.* **183** 22
Kudar J 1926 *Ann. Phys.* **81** 632
Gordon W 1926 *Z. Phys.* **40** 117
- [4] The notation used in this series of papers is the following. $g^{\mu\nu} = \text{diag}(1, -1, -1, -1)$, $\hbar = 1$, $c = 1$. $p_\mu = (E, \mathbf{p})$ is the canonical 4-momentum, while $\pi_\mu = (\pi, \boldsymbol{\pi})$ is the mechanical 4-momentum. $A_\mu = (A_0, \mathbf{A})$ is the electromagnetic 4-potential, $\pi_\mu = p_\mu - eA_\mu$. $\partial_\mu = (\partial_0, \boldsymbol{\partial}) = (\frac{\partial}{\partial t}, -\nabla)$. $p_\mu \rightarrow i\partial_\mu$, $\pi_\mu \rightarrow iD_\mu$ is the quantization process. $D_\mu = \partial_\mu + (ie)A_\mu$ is the minimal coupling. $\mathbf{D} = \boldsymbol{\partial} + ie\mathbf{A}$. $a \cdot b = a_\mu b^\mu$ is the Lorentz scalar product. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, $\mathbf{E} = -\nabla A_0 - (\partial\mathbf{A}/\partial t)$, $\mathbf{B} = \nabla \times \mathbf{A}$. The gamma matrices are defined as follows. $\gamma_\mu = (\gamma_0, \boldsymbol{\gamma})$, $(\gamma_0 = \gamma^0)$. In the standard representation,

$$\gamma_{\text{STD}}^0 = \begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} \quad \boldsymbol{\gamma}_{\text{STD}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & \mathbf{0} \end{pmatrix} \quad \gamma_{\text{SSTD}} = \begin{pmatrix} \mathbf{0} & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} \end{pmatrix}$$

$$\gamma_{\text{Weyl}}^\mu = U\gamma_{\text{STD}}^\mu U^\dagger \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{1}_2 & \mathbf{1}_2 \\ \mathbf{1}_2 & -\mathbf{1}_2 \end{pmatrix}$$

$\boldsymbol{\sigma}$ and $\boldsymbol{\tau}_i$ are both the conventional Pauli matrices and $\not{D} = \gamma^\mu D_\mu$.

- [5] Dirac P A M 1928 *Proc. R. Soc. (London) A* **117** 610
- [6] Klein O and Nishina Y 1929 *Z. Phys.* **52** 853
- [7] Waller I 1930 *Z. Phys.* **61** 837
Tamm I 1930 *Z. Phys.* **62** 545
- [8] Anderson C A 1932 *Science* **76** 238
- [9] Feshbach H and Villars F 1958 *Rev. Mod. Phys.* **30** 24. Equation (2.15) in their paper is what we denote as the two-component spin-0 equation. Feshbach and Villars clearly state that their work is based upon earlier material, in particular Taketani M and Sakata S 1940 *Proc. Math. Phys. Soc. Japan* **22** 757, and Heitler W 1943 *Proc. R. Irish Acad.* **49** 1. Taketani and Sakata were the first to write Kemmer's (Kemmer N 1939 *Proc. R. Soc. (London) A* **173** 91) ten-component first-order equation for spin-1 particles as a six-component equation in Hamiltonian form. The two-component equation for (free) spin-0 particles is

easily deduced from the six-component equation, as shown by Heitler. Heitler also gives an extensive discussion of the equation, much of which is used by Feshbach–Villars in their paper. Feshbach and Villars, however, were the first to write the equation with a minimal coupling, and also to derive the equation directly from the Klein–Gordon equation by linearizing the time derivative.

- [10] Feynman R P and Gell-Mann M 1958 *Phys. Rev.* **109** 193, equation (3)
- [11] Staudte D S 1996 *J. Phys. A: Math. Gen.* **29** 169
- [12] Bethe H A and Salpeter E E 1977 *Quantum Mechanics of One and Two Electron Atoms* (New York: Plenum) p 63
- [13] Messiah A 1981 *Quantum Mechanics* vol 2 (Amsterdam: North-Holland) p 930
- [14] Auvil P R and Brown L M 1975 *Am. J. Phys.* **46** 679
- [15] We use the convention of Merzbacher E 1970 *Quantum Mechanics* (New York: Wiley) equation (16.80)