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

1996 J. Phys. A: Math. Gen. 29 169

(http://iopscience.iop.org/0305-4470/29/1/018)

J. Phys. A: Math. Gen. **29** (1996) 169–192. Printed in the UK

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

D S Staudte

Department of Theoretical Physics, Research School of Physical Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia and Fysik 3A, Chalmers Tekniska Högskola, 412 96 Göteborg, Sweden

Received 2 May 1995

Abstract. The eight-component relativistic wave equation for spin-$\frac{1}{2}$ particles derived in the preceding paper is studied from a mathematical viewpoint. It is shown that it is relativistically covariant, albeit not manifestly so. It has an enlarged solution space when compared to the Dirac equation. A derivation of the equation in an arbitrary gamma matrix representation is presented. Kronecker products are used to display the analogy between the eight-component equation and the two-component spin-0 equation. The linearization procedure used to obtain the eight-component equation from the second-order spin-$\frac{1}{2}$ equation is found to have a natural mathematical validation in the study of indefinite inner product spaces. The (relativistic) quantum mechanical formalism for the eight-component equation is constructed. This features the use of an indefinite inner product and some results are presented to show how the usual quantum mechanical formalism is generalized to account for this. The reason for the decoupling of the equation into two four-component equations in the Weyl representation of the gamma matrices is given. It is shown that only one of the decoupled parts has to be solved for any problem of a single particle coupled to an external electromagnetic field. Comments are made on the solution of the hydrogen atom presented in the preceding paper.

1. Introduction

This is the second of a series of papers concerning the development of a spin-$\frac{1}{2}$ relativistic wave equation which involves the use of an indefinite inner product in the description of its solution space. In the first paper [1] the equation, which has eight components and is in Hamiltonian form, was derived. It was used to produce the relativistic bound-state energy eigenvalue spectra and wavefunctions for hydrogenic atoms to show that it could be used for a single particle minimally coupled to a classical external electromagnetic field.

This paper seeks to establish, from a mathematical viewpoint, the validity of the equation in relativistic quantum mechanics.

Historically, the Dirac equation [2] was constructed in order to incorporate the effects of special relativity into the formalism of non-relativistic quantum mechanics. A relativistic wave equation in Hamiltonian form was sought which retained the probability interpretation of quantum mechanics. The Dirac equation proved to be such an equation although the negative energy solutions required further interpretation. The relativistic nature of the Dirac equation can easily be seen by considering the representation theory of the Lorentz group. The probability interpretation is one result of the mathematical formalism of quantum mechanics, based upon the use of a Hilbert space.

In this paper, the eight-component equation is discussed using elements of the above mentioned mathematics. It will be shown that, despite its appearance, the equation is

0305-4470/96/010169+24$19.50

relativistically covariant. It requires the use of an indefinite inner product, which is one of two major distinguishing factors from the formalism associated with the Dirac equation. The other is that the equation has an enlarged solution space, the dimension of which is twice that of the Dirac solution space. It is found that there are two advantages in using the equation as opposed to the second-order spin-$\frac{1}{2}$ equation. These are its Hamiltonian form, which allows the possibility of developing a quantum mechanical formalism, and a very natural interpretation using considerations from the theory of indefinite inner product spaces. A number of results are given which provide the basis for the required quantum mechanical formalism associated to the equation.

In section 2 a brief summary of the results of the preceding paper is given. The eight-component equation is hereafter referred to as the FV$\frac{1}{2}$ equation due to its similar form to the two-component spin-0 relativistic wave equation which we call the Feshbach–Villars equation [3]. The FV$\frac{1}{2}$ equation is obtained from a spin-$\frac{1}{2}$ relativistic wave equation containing second-order derivatives in both time and space [4] (which we refer to as the KG$\frac{1}{2}$ equation due to its similar form to the spin-0 Klein–Gordon equation) by linearizing only the time derivative. There is a one-to-one correspondence between solutions of the FV$\frac{1}{2}$ and KG$\frac{1}{2}$ equations and the equations themselves are equivalent. Hence, by studying the KG$\frac{1}{2}$ equation one can obtain direct information about the FV$\frac{1}{2}$ equation. The KG$\frac{1}{2}$ and Dirac equations are in section 3 derived using spinors and their product representations from which their relativistic covariance is easily seen. Thus the FV$\frac{1}{2}$ equation is also covariant if not manifestly so. The use of spinor quantities illustrates the mathematical origins of the Dirac and KG$\frac{1}{2}$ equations.

In section 4 it is shown that the solution space of the KG$\frac{1}{2}$ equation is twice the dimension of the solution space of the Dirac equation.

A comparison of the FV$\frac{1}{2}$ equation with the Feshbach–Villars equation (hereafter referred to as the FV0 equation) is given in section 5, followed by a derivation of the FV$\frac{1}{2}$ equation in an arbitrary gamma matrix representation. The linearization procedure used to derive the FV$\frac{1}{2}$ equation is then justified in section 6 from a consideration of conserved currents for Klein–Gordon type equations in terms of inner product spaces. It is shown that they can be written most simply and transparently by linearizing only the time derivative in the Klein–Gordon type equations. This provides a rationale for the process beyond that in [3], which is applicable to both the spin-0 and spin-$\frac{1}{2}$ cases. The FV$\frac{1}{2}$ equation is written using Kronecker products, which simplifies its interpretation and preserves the gamma matrix algebra, which is present in any relativistic spin-$\frac{1}{2}$ equation, in a manifest form.

A useful result of the FV$\frac{1}{2}$ equation is that it decouples into two separate equations in the Weyl representation of the gamma matrices. That this should happen is shown in section 7 using the theory of projectors. A method is given in which only one of the decoupled equations needs to be solved for problems of a single particle minimally coupled to a classical external electromagnetic field and how the full eight-component wavefunctions can easily be constructed from this.

The theory of indefinite inner product spaces is used in section 8 to construct the quantum mechanical formalism associated to the FV$\frac{1}{2}$ equation. The Hamiltonian is shown to be pseudo-Hermitian and thus plays the same rôle as the ordinary quantum mechanical Hamiltonian. The definition of the expectation value of an operator as given in [3] is clarified. The pseudo-unitary transformations in [3] are extended to the spin-$\frac{1}{2}$ case and Kronecker products used to show that transformations between gamma matrix representations take, as they must, the usual unitary form.

In section 9, the indefinite inner product is written in a form which takes advantage of the decoupling of the FV$\frac{1}{2}$ equation.

In the preceding paper the equation was applied to hydrogenic atoms. It was seen that the bound-state energy eigenvalue spectra obtained are identical to those from the Dirac equation, but that the wavefunctions differ. In section 10 it is shown that the angular wavefunctions are identical, retaining the angular momentum classification of states. The exact relationship between the radial wavefunctions is found. The FV$\frac{1}{2}$ wavefunctions are consistent with the literature, suggesting that the solution method given in [1] and section 7 of this paper is correct. These results hold for those bound state solutions constructed so that $\langle\Psi|\Psi\rangle = +1$. The other solutions will be discussed in conjunction with the physical interpretation of the enlarged solution space using the indefinite inner product in a later paper.

Finally, in section 11 some conclusions are presented together with an outline of forthcoming papers.

2. The FV$\frac{1}{2}$ equation

In the preceding paper [1] the FV$\frac{1}{2}$ equation was derived. The FV0 equation is obtained from the Klein–Gordon equation (KG0 equation) by linearizing the time derivative. It is in Hamiltonian form, $\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_2\Psi_{\mathrm{FV0}} = H_{\mathrm{FV0}}\Psi_{\mathrm{FV0}}$, with [3]

$$H_{\mathrm{FV0}} = -\frac{\boldsymbol{D}^2}{2m}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} + m\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + eA_0\mathbf{1}_2 \tag{1}$$

$$\Psi_{\mathrm{FV0}} = \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \Psi_{\mathrm{KG0}} + m^{-1}\mathrm{i}D_0\Psi_{\mathrm{KG0}} \\ \Psi_{\mathrm{KG0}} - m^{-1}\mathrm{i}D_0\Psi_{\mathrm{KG0}} \end{pmatrix}. \tag{2}$$

The FV$\frac{1}{2}$ equation was derived via an analogous linearization of the KG$\frac{1}{2}$ equation. The Weyl representation of the gamma matrices was chosen, where the KG$\frac{1}{2}$ equation decouples into two two-component equations. Upon linearization two four-component equations were obtained, these are $\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4\Psi_\xi = H_\xi\Psi_\xi$ and $\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4\Psi_{\dot\eta} = H_{\dot\eta}\Psi_{\dot\eta}$ with

$$H_\xi = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \otimes \frac{1}{2m}(-\boldsymbol{D}^2\mathbf{1}_2 + \mathrm{i}e\boldsymbol{\sigma}\cdot(\boldsymbol{E} + \mathrm{i}\boldsymbol{B})) + m\begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} + eA_0\mathbf{1}_4 \tag{3}$$

$$H_{\dot\eta} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \otimes \frac{1}{2m}(-\boldsymbol{D}^2\mathbf{1}_2 - \mathrm{i}e\boldsymbol{\sigma}\cdot(\boldsymbol{E} - \mathrm{i}\boldsymbol{B})) + m\begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} + eA_0\mathbf{1}_4. \tag{4}$$

These two equations combine into an eight-component equation, which is (one form of) the FV$\frac{1}{2}$ equation

$$\begin{pmatrix} (\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4 - H_\xi) & 0 \\ 0 & (\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4 - H_{\dot\eta}) \end{pmatrix}\begin{pmatrix} \Psi_\xi \\ \Psi_{\dot\eta} \end{pmatrix} = 0. \tag{5}$$

The equation $\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4\Psi_\xi = H_\xi\Psi_\xi$ was solved exactly in the presence of a minimally coupled Coulomb potential to produce the relativistic bound-state energy eigenvalue spectra and wavefunctions for hydrogenic atoms. The spectra are the same as those obtained using the Dirac equation, but the wavefunctions deviate slightly due to the fact that a different Hamiltonian is used.

3. The KG$\frac{1}{2}$ and Dirac equations in spinor notation

An important step in the development of the FV$\frac{1}{2}$ equation is to show that it is indeed relativistic. Since the FV$\frac{1}{2}$ equation is equivalent to the KG$\frac{1}{2}$ equation, it suffices to show the covariance of the KG$\frac{1}{2}$ equation. This is most easily done using spinor representations of the Lorentz Group. The KG$\frac{1}{2}$ equation is in conventional 4-vector notation $((\mathrm{i}\not{\partial})^2 - m^2\mathbf{1}_4)\Psi_{\mathrm{KG}1/2} = 0$, whereas the Dirac equation is $(\mathrm{i}\not{\partial} - m\mathbf{1}_4)\Psi_{\mathrm{D}} = 0$.

These equations will now be derived using quantities constructed from the spinor irreducible representations of the proper (inhomogeneous) Lorentz group, hereafter referred to as the PLG. The PLG is defined to contain rotations in three-dimensional space plus Lorentz boosts. The two spinor irreducible representations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ of the PLG are two-component quantities and here they are labelled ξ^α and $\eta_{\dot\beta}$ [5]. The spinor equivalent of a 4-vector is constructed from the product of two separate spinors ξ^α and $\Theta^{\dot\beta} = (\Theta^\beta)^*$, and is written $V^{\alpha\dot\beta}$. Written as a 2×2 matrix

$$[V^{\alpha\dot\beta}] = \begin{pmatrix} V^{0\dot0} & V^{0\dot1} \\ V^{1\dot0} & V^{1\dot1} \end{pmatrix} = v_0\mathbf{1}_2 + \boldsymbol{\sigma}\cdot\boldsymbol{v} \tag{6}$$

where $v_\mu = (v_0, \boldsymbol{v})$ is a covariant 4-vector. $V_{\dot\alpha\beta}$ is obtained from $V^{\alpha\dot\beta}$ by spatial inversion. $P^{\alpha\dot\beta}$ is used to denote the quantity related to $\mathrm{i}\partial_\mu$ by the above prescription. A minimal coupling to a classical external electromagnetic field is introduced by the replacement

$$P^{\alpha\dot\beta} \to \Pi^{\alpha\dot\beta} = P^{\alpha\dot\beta} - eA^{\alpha\dot\beta} \tag{7}$$

which is the spinor equivalent of the minimal coupling

$$\partial_\mu \to D_\mu = \partial_\mu + \mathrm{i}eA_\mu. \tag{8}$$

Consider a first-order partial differential equation acting on the two irreducible representations ξ^α and $\eta_{\dot\beta}$

$$\Pi^{\alpha\dot\beta}\eta_{\dot\beta} = m\xi^\alpha \tag{9}$$

$$\Pi_{\dot\beta\alpha}\xi^\alpha = m\eta_{\dot\beta}. \tag{10}$$

These are the simplest first-order manifestly covariant equations one can write. If equations (9) and (10) are combined into a single four-component equation, one obtains the Dirac equation in the Weyl (also known as the spinor [5] or helicity) representation of the gamma matrices. Here, the gamma matrices take the form [6]

$$\gamma^0 = \begin{pmatrix} \mathbf{0} & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} \end{pmatrix} \qquad \gamma = \begin{pmatrix} \mathbf{0} & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix} \qquad \gamma_5 = \begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix}. \tag{11}$$

The simplest manifestly covariant second-order equations are

$$\Pi^{\alpha\dot\beta}\Pi_{\dot\beta\gamma}\xi^\gamma = m^2\xi^\alpha \tag{12}$$

$$\Pi_{\dot\alpha\beta}\Pi^{\beta\dot\gamma}\eta_{\dot\gamma} = m^2\eta_{\dot\alpha}. \tag{13}$$

There are two separate uncoupled equations, which transform into each other under spatial inversion. Each equation is a second-order two-component equation. If one wants to use these equations in relativistic quantum mechanics, (12) and (13) should be combined into one four-component equation, as then the whole equation will be invariant under spatial inversion as required for electromagnetic interactions. This combined four-component equation is the $\mathrm{KG}\frac{1}{2}$ equation in the Weyl representation of the gamma matrices. Equations (12) and (13) themselves are identically the equations (14), [1] for $\Psi_{\pm}$. The choice of m^2 for the constants in equations (12) and (13) is justified as for free particles

$$\Pi^{\alpha\dot{\beta}}\Pi_{\dot{\beta}\gamma} = P^{\alpha\dot{\beta}}P_{\dot{\beta}\gamma} = -\partial^{\mu}\partial_{\mu}\delta^{\alpha}{}_{\gamma} \tag{14}$$

and one obtains the free particle Klein–Gordon equations for the two-component spinors ξ^{α} and $\eta_{\dot{\beta}}$, $(\partial^{\mu}\partial_{\mu}+m^2)\mathbf{1}_2\xi^{\alpha}=0$ and $(\partial^{\mu}\partial_{\mu}+m^2)\mathbf{1}_2\eta_{\dot{\beta}}=0$. Using this notation it can be seen that while the $\mathrm{KG}\frac{1}{2}$ equation decouples into two equations, it is not possible to construct a two-component first-order equation due to the fact that $\Pi^{\alpha\dot{\beta}}$ couples the two types of spinor irreducible representations of the PLG. The decoupling of the $\mathrm{KG}\frac{1}{2}$ equation occurs only in the Weyl representation, where $\Psi_{\mathrm{KG}1/2}=(\xi^{\alpha},\eta_{\dot{\beta}})^{T}$.

The use of spinor quantities provides a natural way to derive spin-$\frac{1}{2}$ relativistic wave equations of either first or second order with the correct minimal coupling. The covariance of these equations is automatic and manifest. The method of construction via spinors shows that the $\mathrm{KG}\frac{1}{2}$ equation can be derived *a priori*, rather than only by 'squaring' the Dirac equation. It is also observed that the $\mathrm{FV}\frac{1}{2}$ equation originates, not from the Dirac equation, but from another, equally mathematically valid, relativistic wave equation. The covariance of the $\mathrm{KG}\frac{1}{2}$ equation establishes the covariance of the $\mathrm{FV}\frac{1}{2}$ equation. Also, the fact that the $\mathrm{KG}\frac{1}{2}$ equation describes spin-$\frac{1}{2}$ particles means that the $\mathrm{FV}\frac{1}{2}$ equation does likewise.

4. The solution spaces of the $\mathrm{KG}\frac{1}{2}$ and Dirac equations

The Dirac equation in conventional notation is

$$(\mathrm{i}\not{\partial}-m\mathbf{1}_4)\Psi=0 \qquad \text{or} \qquad \mathrm{i}\not{\partial}\Psi=m\mathbf{1}_4\Psi. \tag{15}$$

'Squaring' (15) gives

$$(\mathrm{i}\not{\partial})^2\Psi=m^2\mathbf{1}_4\Psi \qquad \text{or} \qquad ((\mathrm{i}\not{\partial})^2-m^2\mathbf{1}_4)\Psi=0 \tag{16}$$

which is the $\mathrm{KG}\frac{1}{2}$ equation. It can be seen that if one started with equation (15), replaced m by $-m$, and then 'squared' the equation, that equation (16) is again obtained. Let $\mathcal{D}_{\pm}$ be the solution spaces of the equations

$$(\mathrm{i}\not{\partial}\pm m\mathbf{1}_4)\Psi_{\pm}=0 \tag{17}$$

and $\mathcal{F}$ be the solution space of equation (16). Some relationships between the solution spaces of equations (17) and (16) are listed below [7]

$$\mathcal{D}_{+}\cup\mathcal{D}_{-}\subset\mathcal{F} \tag{18}$$

$$\forall\Psi_{-}\in\mathcal{D}_{-}\ \exists|\Psi_{+}\in\mathcal{D}_{+}\ni\Psi_{+}=\gamma_5\Psi_{-} \tag{19}$$

$$\forall\Psi\in\mathcal{F}\ \exists|\Psi_{+}\in\mathcal{D}_{+}\wedge\Psi_{-}\in\mathcal{D}_{-}\ni\Psi=\Psi_{-}-\Psi_{+},\ \Psi_{\pm}=(2m)^{-1}(\mathrm{i}\not{\partial}\mp m\mathbf{1}_4)\Psi \tag{20}$$

$$\mathcal{D}_{+}\cap\mathcal{D}_{-}=\{\mathbf{0}\}. \tag{21}$$

Since $\mathcal{D}_+$ and $\mathcal{D}_-$ are disjoint (equation (21)), and the dimension of $\mathcal{D}_-$ is equal to the dimension of $\mathcal{D}_+$ (which follows from equation (19) since γ_5 is non-singular), then $\mathcal{D}_- \cup \mathcal{D}_+$ has twice the dimension of $\mathcal{D}_-$. Thus, according to equation (18), the solution space $\mathcal{F}$ has dimension at least twice that of the Dirac equation solution space. It is exactly twice because equation (20) shows that $\mathcal{D}_- \cup \mathcal{D}_+$ spans $\mathcal{F}$.

The enlarged solution space of the KG$\frac{1}{2}$ equation when compared to the Dirac equation is a major distinguishing factor between these equations. The one-to-one correspondence between solutions of the FV$\frac{1}{2}$ and KG$\frac{1}{2}$ equations shows that the solution space of the FV$\frac{1}{2}$ equation is twice the dimension of the Dirac equation solution space.

Equation (18) shows that in general, rather than a Dirac solution, a linear combination of a Dirac and a separate solution is to be considered as a solution of the KG$\frac{1}{2}$ equation. Equation (19) shows that there is a one-to-one correspondence between solutions in the spaces $\mathcal{D}_-$ and $\mathcal{D}_+$. However, equation (21) emphasizes that any solution or linear combination of solutions in $\mathcal{D}_-$ is linearly independent to each and every solution in $\mathcal{D}_+$. Finally, equation (20) shows how any solution of the KG$\frac{1}{2}$ equation can be written as a linear combination of elements of $\mathcal{D}_-$ and $\mathcal{D}_+$. The solutions Ψ_- and Ψ_+ in equation (20) do not, however, satisfy equation (19).

5. The FV$\frac{1}{2}$ equation in an arbitrary gamma matrix representation

In the preceding paper, the FV$\frac{1}{2}$ equation was derived in the Weyl representation of the gamma matrices. This has the advantage that the FV$\frac{1}{2}$ equation decouples into two separate four-component equations in Hamiltonian form, with Hamiltonians given by equations (3) and (4). The Hamiltonians (3) and (4) differ from H_{FV0} (1) in that they each contain twice as many components as H_{FV0} and also in that they contain a term representing the interaction of the spin with the external field. The external electromagnetic field appears in the well known [5] relativistic combinations $(\boldsymbol{E} + \mathrm{i}\boldsymbol{B})$ and $-(\boldsymbol{E} - \mathrm{i}\boldsymbol{B})$. Under spatial inversion these complex 3-vectors interchange. These combinations are already manifest in the second-order equations (14), [1].

The non-relativistic generalization of the spin-0 Schrödinger equation to the spin-$\frac{1}{2}$ Pauli equation [8] involves the substitution

$$\boldsymbol{D}^2 \to (\boldsymbol{\sigma} \cdot \boldsymbol{D})^2 = \boldsymbol{D}^2 \mathbf{1}_2 + e\boldsymbol{\sigma} \cdot \boldsymbol{B}. \tag{22}$$

The relativistic generalization contains not just the magnetic field, but specific combinations of the magnetic and electric fields. The wavefunctions that the Hamiltonians (3) and (4) act on are of similar form to (2) except that each component is multiplied by a two-component spinor as follows. The two equations (12) and (13) each have a two-component wavefunction (now) written as ξ and $\dot{\eta}$ respectively. Combined together, the wavefunction becomes $\Psi_{\mathrm{KG1/2}} = (\xi, \dot{\eta})^T$. In order to obtain the FV$\frac{1}{2}$ wavefunction from this four-component wavefunction, define

$$\Psi_{\mathrm{KG1/2}} = \begin{pmatrix} \xi \\ \dot{\eta} \end{pmatrix} = \psi \begin{pmatrix} \xi_0 \\ \dot{\eta}_0 \end{pmatrix} \tag{23}$$

where ψ is a scalar containing everything in the wavefunction invariant under spatial inversion. ξ_0 and $\dot{\eta}_0$ each have two components and contain only the parts that transform into

each other under spatial inversion. Any time dependence of the wavefunction is contained in ψ as this will not be altered under spatial inversion. Using (23) one can again define

$$\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \psi + m^{-1} \mathrm{i} D_0 \psi \\ \psi - m^{-1} \mathrm{i} D_0 \psi \end{pmatrix}. \tag{24}$$

This definition is analogous to (2). The wavefunctions corresponding to (3) and (4) become respectively

$$\Psi_\xi = \begin{pmatrix} \phi \xi_0 \\ \chi \xi_0 \end{pmatrix} \qquad \text{and} \qquad \Psi_{\dot\eta} = \begin{pmatrix} \phi \dot\eta_0 \\ \chi \dot\eta_0 \end{pmatrix}. \tag{25}$$

The eight-component wavefunction in equation (5) is simply $\Psi_{\mathrm{FV1/2}} = (\Psi \otimes \xi_0, \Psi \otimes \dot\eta_0)^T$ and equation (5) itself can be written as

$$\mathrm{i} \tfrac{\partial}{\partial t} \mathbf{1}_8 \Psi_{\mathrm{FV1/2}} = H_{\mathrm{FV1/2}} \Psi_{\mathrm{FV1/2}} \qquad H_{\mathrm{FV1/2}} = \begin{pmatrix} H_\xi & 0 \\ 0 & H_{\dot\eta} \end{pmatrix} \tag{26}$$

$$H_\xi = (\tau_3 + \mathrm{i}\tau_2) \otimes \left(\frac{1}{2m} (-\boldsymbol{D}^2 \mathbf{1}_2 + \mathrm{i} e \boldsymbol{\sigma} \cdot (\boldsymbol{E} + \mathrm{i}\boldsymbol{B})) \right) + m(\tau_3 \otimes \mathbf{1}_2) + e A_0 (\mathbf{1}_2 \otimes \mathbf{1}_2) \tag{27a}$$

$$H_{\dot\eta} = (\tau_3 + \mathrm{i}\tau_2) \otimes \left(\frac{1}{2m} (-\boldsymbol{D}^2 \mathbf{1}_2 - \mathrm{i} e \boldsymbol{\sigma} \cdot (\boldsymbol{E} - \mathrm{i}\boldsymbol{B})) \right) + m(\tau_3 \otimes \mathbf{1}_2) + e A_0 (\mathbf{1}_2 \otimes \mathbf{1}_2) \tag{27b}$$

where the τ_i are the usual Pauli matrices.

The FV$\frac{1}{2}$ equation will now be derived for an arbitrary gamma matrix representation. Consider the KG$\frac{1}{2}$ equation and define $\Psi_{\mathrm{KG1/2}}$ in an arbitrary representation by $\Psi_{\mathrm{KG1/2}} = (\alpha, \beta)^T$, where

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = U \begin{pmatrix} \xi \\ \dot\eta \end{pmatrix} \tag{28}$$

and the gamma matrices in this representation are $\gamma^\mu_U = U \gamma^\mu_{\mathrm{Weyl}} U^\dagger$. In analogy to equation (23), define

$$\Psi_{\mathrm{KG1/2}} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \psi \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \tag{29}$$

where ψ is identical to the ψ defined by equation (23), as under a unitary transformation of the gamma matrices ψ will remain unchanged.

The wavefunction $\Psi_{\mathrm{FV1/2}}$ in the Weyl representation is given in equation (5). However, for an arbitrary gamma matrix representation, the wavefunction $\Psi_{\mathrm{FV1/2}} = (\Psi \otimes \alpha_0, \Psi \otimes \beta_0)^T$ is not convenient to display the analogous nature of the FV$\frac{1}{2}$ and FV0 equations. The purpose of defining ψ via equation (23) is to obtain a wavefunction of the form $(\phi, \chi)^T$ multiplied by a spinor wavefunction, where $(\phi, \chi)^T$ plays as close a rôle as possible to $(\phi, \chi)^T$ of the FV0 equation. This makes it possible to write the equation, a conserved current, operators and expectation values in a way analogous to the FV0 equation, and hence to develop an analogous relativistic quantum mechanics for the FV$\frac{1}{2}$ equation. We instead define $\Psi'_{\mathrm{FV1/2}} = \Psi \otimes \Psi_0$ where $\Psi_0 = (\alpha_0, \beta_0)^T$. The linearization procedure involves rewriting $\Psi_{\mathrm{KG1/2}} = \psi\Psi_0$ ($= \psi \otimes \Psi_0$ as ψ is a single component object) as $\Psi'_{\mathrm{FV1/2}} = \Psi \otimes \Psi_0$.

Note that the bispinor Ψ_0 remains unchanged by the linearization process. $\Psi'_{\mathrm{FV}1/2}$ satisfies an eight-component equation $\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_8\Psi'_{\mathrm{FV}1/2} = H'_{\mathrm{FV}1/2}\Psi'_{\mathrm{FV}1/2}$, with

$$H'_{\mathrm{FV}1/2} = (\tau_3 + \mathrm{i}\tau_2) \otimes \left(\frac{1}{2m}\left(-\boldsymbol{D}^2\mathbf{1}_4 + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\right) + m(\tau_3 \otimes \mathbf{1}_4) + eA_0(\mathbf{1}_2 \otimes \mathbf{1}_4). \tag{30}$$

$H_{\mathrm{FV}0}$ is given here for comparison

$$H_{\mathrm{FV}0} = (\tau_3 + \mathrm{i}\tau_2)\left(\frac{1}{2m}(-\boldsymbol{D}^2)\right) + m\tau_3 + eA_0\mathbf{1}_2. \tag{31}$$

The only changes to the $\mathrm{FV}\frac{1}{2}$ equation under a unitary transformation of the gamma matrices are to change the $\sigma^{\mu\nu}$ term in the Hamiltonian (30) and the bispinor Ψ_0 in the wavefunction. The $\sigma^{\mu\nu}$ term is the only one that couples α_0 and β_0, and in the Weyl representation $\sigma^{\mu\nu}$ is block diagonal which allows separate equations for α_0 and β_0 to be written. These separate equations have four components with wavefunctions $\Psi \otimes \alpha_0 = \Psi_{\xi}$ and $\Psi \otimes \beta_0 = \Psi_{\dot{\eta}}$. The analogous nature of $H'_{\mathrm{FV}1/2}$ and $H_{\mathrm{FV}0}$ is clear. Throughout this paper, $\Psi_{\mathrm{FV}1/2}$ and $H_{\mathrm{FV}1/2}$ refer to the wavefunction and Hamiltonian written as in equations (3–5) while $\Psi'_{\mathrm{FV}1/2}$ and $H'_{\mathrm{FV}1/2}$ refer to the wavefunction $\Psi \otimes \Psi_0$ and the Hamiltonian given by equation (30).

The $\mathrm{KG}\frac{1}{2}$ equation (16) can be conveniently written as

$$\left((D^\mu D_\mu + m^2)\mathbf{1}_4 + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\Psi_{\mathrm{KG}1/2}(\boldsymbol{x}) = 0 \tag{32}$$

which shows that the spin term is preserved in the Hamiltonian (30).

6. The Feshbach–Villars linearization procedure

The linearization procedure used to obtain the $\mathrm{FV}\frac{1}{2}$ equation in any of its forms from the $\mathrm{KG}\frac{1}{2}$ equation is analogous to that used by Feshbach and Villars [3] in the spin-0 case and for convenience will be referred to as the Feshbach–Villars linearization procedure regardless of whether it is applied in the spin-$\frac{1}{2}$ or spin-0 case. It involves the rewriting of a manifestly covariant, second order in the time derivative equation as a non-manifestly covariant, first order in the time derivative equation . Why should this be useful? To answer this question, some elements of the formalism of non-relativistic quantum mechanics are first reviewed.

In non-relativistic quantum mechanics a system is represented by a vector $|\psi\rangle \in \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space [9]. In the position representation, the wavefunction $\psi(\boldsymbol{x})$ is a complex valued function of the real variable $\boldsymbol{x}$ and satisfies an equation in Hamiltonian form (the Schrödinger equation), $\mathrm{i}\frac{\partial}{\partial t}\psi(\boldsymbol{x}) = H(\boldsymbol{x})\psi(\boldsymbol{x})$, where $H(\boldsymbol{x})$ is the Hamiltonian given by

$$H(\boldsymbol{x}) = \frac{-\boldsymbol{D}^2}{2m} + eA_0. \tag{33}$$

This equation is obtained from the non-relativistic energy–momentum relation, $E = p^2/2m$, by the replacement $p_\mu \to \mathrm{i}\partial_\mu$ together with a minimal coupling. To describe a system at any time t in the future, one needs to know only H and $|\psi(t_0)\rangle$. The probability interpretation of a state $|\psi\rangle$ representing a single particle is given by $\langle\psi|\psi\rangle = +1$, which in

the position representation becomes $\langle\psi|\psi\rangle = \int j^0 \,\mathrm{d}^3\boldsymbol{x} = +1$, where $j^0 = \psi^*(\boldsymbol{x})\psi(\boldsymbol{x})$. An operator Ω is Hermitian if $\Omega^\dagger = (\Omega^*)^T = \Omega$, and a transformation U unitary if $U^\dagger = U^{-1}$. Hermitian operators have real eigenvalues and expectation values, while unitary transformations preserve the value of $\langle\psi|\psi\rangle$.

The definitions given above of $|\psi\rangle$, $\psi(\boldsymbol{x})$, $\langle\psi|\psi\rangle$, $\langle\Omega\rangle$, $\Omega^\dagger$, etc follow from the mathematics of inner product spaces [9, 10]. A vector space V with an inner product defined on it is known as an inner product space. If $\boldsymbol{w}, \boldsymbol{y} \in V$, then $\langle\boldsymbol{w}|\boldsymbol{y}\rangle = \langle\boldsymbol{y}|\boldsymbol{w}\rangle^*$ for vectors $\boldsymbol{w}$ and $\boldsymbol{y}$ defined on the field of complex numbers. An operator Ω acts on a vector $\boldsymbol{y}$ to give $\Omega\boldsymbol{y} \in V$. In general, $\langle\boldsymbol{w}|\Omega\boldsymbol{y}\rangle = \langle\Omega^\sharp\boldsymbol{w}|\boldsymbol{y}\rangle$ for some $\Omega^\sharp$. If $\Omega^\sharp = \Omega$, then Ω is Hermitian. If $\boldsymbol{y} \to \boldsymbol{y}' = U\boldsymbol{y}, \forall\boldsymbol{y} \in V$, then $\langle\boldsymbol{w}'|\boldsymbol{y}'\rangle = \langle U\boldsymbol{w}|U\boldsymbol{y}\rangle = \langle\boldsymbol{w}|\boldsymbol{y}\rangle$ if $U^\sharp = U^{-1}$. U is unitary if $U^\sharp = U^{-1}$.

A Hilbert space $\mathcal{H}$ is an infinite dimensional inner product space. The space $L^2(-\infty, +\infty)$ with representatives $f(x)$ which are square integrable functions of some real variable x is a Hilbert space and is the model for quantum mechanical Hilbert spaces. With the real variable x being $\boldsymbol{x}$, the inner product is

$$\langle\boldsymbol{w}|\boldsymbol{y}\rangle = \int \boldsymbol{w}^\dagger(\boldsymbol{x})\boldsymbol{y}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}. \tag{34}$$

A general finite dimensional inner product space is defined by $\langle\boldsymbol{w}|\boldsymbol{y}\rangle = \boldsymbol{w}^\dagger\boldsymbol{K}\boldsymbol{y}$, with $\boldsymbol{K}$ a Hermitian, non-singular matrix. The space is proper if $\boldsymbol{K}$ is positive definite ($\boldsymbol{w}^\dagger\boldsymbol{K}\boldsymbol{w} \geq 0\ \forall\boldsymbol{w}$, with equality only for $\boldsymbol{w} = \boldsymbol{0}$). It is improper (indefinite) if $\boldsymbol{K}$ is Hermitian and non-singular but the quantity $\boldsymbol{w}^\dagger\boldsymbol{K}\boldsymbol{w}$ is indefinite in sign. A more general form of equation (34) is given by

$$\langle\boldsymbol{w}|\boldsymbol{y}\rangle = \int \boldsymbol{w}^\dagger(\boldsymbol{x})\boldsymbol{K}\boldsymbol{y}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}. \tag{35}$$

Relativistic wave equations are based on the energy–momentum relation $p^\mu p_\mu = m^2$. The first published relativistic wave equation was the KG0 equation

$$(D^\mu D_\mu + m^2)\Psi_{\mathrm{KG0}}(\boldsymbol{x}) = 0. \tag{36}$$

This equation was intended to be a wave equation analogous to the Schrödinger equation, retaining as much of the non-relativistic quantum mechanical formalism as possible, but using the relativistic energy–momentum relation. Equation (36) is second order in the time derivative and so to specify the time development of the wavefunction requires not only $|\psi(t_0)\rangle$, but also $|\dot{\psi}(t_0)\rangle$. Also, the equation is not in Hamiltonian form. Yet another question appears, what happens to the Hilbert space formalism, is it possible to define an inner product and probability interpretation for equation (36)?

The choice of the Hilbert space $\mathcal{H}$ out of the possible choices of inner product spaces for non-relativistic quantum mechanics can be motivated via the existence of a conserved current density for the Schrödinger equation. The equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\boldsymbol{j} = \frac{\partial j^0}{\partial t} + \nabla\cdot\boldsymbol{j} = \partial_\mu j^\mu = 0 \tag{37}$$

is the well known equation of continuity, with $\rho = j^0$ a density. Such an equation can be derived for the Schrödinger equation $(\mathrm{i}\frac{\partial}{\partial t} - H(\boldsymbol{x}))\psi(\boldsymbol{x}) = 0$ by constructing the following quantity

$$\psi^*(\boldsymbol{x})\left(\mathrm{i}\frac{\partial}{\partial t} - H(\boldsymbol{x})\right)\psi(\boldsymbol{x}) - \left(\left(\mathrm{i}\frac{\partial}{\partial t} - H(\boldsymbol{x})\right)\psi(\boldsymbol{x})^*\right)\psi(\boldsymbol{x}) = 0 \tag{38}$$

which can be written [11] as

$$\frac{\partial}{\partial t}(\psi^*(\boldsymbol{x})\psi(\boldsymbol{x})) + \nabla \cdot \boldsymbol{j} = 0. \tag{39}$$

Hence $\psi^*(\boldsymbol{x})\psi(\boldsymbol{x})$ is a density and its integral over all space $\int \psi^*(\boldsymbol{x})\psi(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}$ is just $\langle\psi|\psi\rangle$, the inner product of two vectors in the quantum mechanical Hilbert space $\mathcal{H}$ using the position representation.

To search for a conserved current density for the KG0 equation a similar procedure is used [12]:

$$\Psi^*_{\mathrm{KG0}}(\boldsymbol{x})((D^\mu D_\mu + m^2)\Psi_{\mathrm{KG0}}(\boldsymbol{x})) - ((D^\mu D_\mu + m^2)\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^*\Psi_{\mathrm{KG0}}(\boldsymbol{x}) = 0. \tag{40}$$

This can also be written in the form of equation (37). j^μ_{KG0} is given by

$$\begin{aligned} j^\mu_{\mathrm{KG0}} &= m^{-1}\mathrm{i}(\Psi^*_{\mathrm{KG0}}(\boldsymbol{x}) \overleftrightarrow{D^\mu} \Psi_{\mathrm{KG0}}(\boldsymbol{x})) \\ &= m^{-1}\mathrm{i}(\Psi^*_{\mathrm{KG0}}(\boldsymbol{x})D^\mu\Psi_{\mathrm{KG0}}(\boldsymbol{x}) - (D^\mu\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^*\Psi_{\mathrm{KG0}}(\boldsymbol{x})). \end{aligned} \tag{41}$$

This method of deriving the conserved current density is analogous to the method of non-relativistic quantum mechanics (equation (38)), even though equation (36) is not in Hamiltonian form. One is seeking to find an equation $\partial_\mu j^\mu = 0$, with j^μ transforming as a 4-vector. j^μ given by equation (41) clearly transforms in the correct manner. This is a derivation from quantum mechanics, not field theory, however, in any case the conserved current derived using the Lagrangian formalism for fields matches the one derived using the formalism of quantum mechanics.

The zeroth component of j^μ given by equation (41) is not positive definite. This was recognized immediately when the KG0 equation was initially published indicating that j^0 is something more than a probability density. This is mentioned in almost every textbook on relativistic quantum mechanics and it is often stated that the equation should also be first order in the time derivative. However, the fact that the Hilbert space formalism is also lost seems to be given less explicit attention. It is not possible to write $\int j^0_{\mathrm{KG0}}\,\mathrm{d}^3\boldsymbol{x}$ in the form $\langle\Psi_{\mathrm{KG0}}|\Psi_{\mathrm{KG0}}\rangle = \int \Psi^\dagger_{\mathrm{KG0}}(\boldsymbol{x})\boldsymbol{K}\Psi_{\mathrm{KG0}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}$ as j^0_{KG0} contains not only the quantity $\Psi_{\mathrm{KG0}}(\boldsymbol{x})$, but also $D^0\Psi_{\mathrm{KG0}}(\boldsymbol{x})$, being

$$j^0_{\mathrm{KG0}} = m^{-1}\mathrm{i}(\Psi^*_{\mathrm{KG0}}(\boldsymbol{x})D^0\Psi_{\mathrm{KG0}}(\boldsymbol{x}) - (D^0\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^*\Psi_{\mathrm{KG0}}(\boldsymbol{x})). \tag{42}$$

Dirac's equation in 1928 [2] was an attempt to regain the 'lost' formalism of quantum mechanics using a relativistic wave equation. It was a linear equation related to the KG0 equation that had the following advantages: (i) it was a first order in the time derivative equation in Hamiltonian form, hence the important time development of the wavefunction was regained; (ii) there existed an inner product formalism based upon a 4-component vector $|\Psi_{\mathrm{D}}\rangle$ in a Hilbert space with

$$\langle\Psi_{\mathrm{D}}|\Psi_{\mathrm{D}}\rangle = \int \Psi^\dagger_{\mathrm{D}}(\boldsymbol{x})\boldsymbol{K}\Psi_{\mathrm{D}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x} \qquad \boldsymbol{K} = \mathbf{1}_4 \tag{43}$$

(iii) the inner product was positive definite which at the time agreed with the interpretations of experiments; (iv) the equation explained the 'spin' of the electron and the hydrogen atom

spectrum. However, the negative energy states signalled the need for a deeper understanding of the situation.

In this work the motivations of Dirac are altered [1] only in that an indefinite inner product similar to that in [3] is desired, together with a relativistic spin-$\frac{1}{2}$ equation that resembles that given in [3]. Given that an equation in Hamiltonian form is sought, some linearization procedure must be applied to the original quadratic relativistic energy–momentum relation $p^\mu p_\mu = m^2$.

Reconsider the zeroth component of the Klein–Gordon current given in equation (42). This involves the product of $(D^0\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^*$ with $\Psi_{\mathrm{KG0}}(\boldsymbol{x})$, and the product of $\Psi^*_{\mathrm{KG0}}(\boldsymbol{x})$ with $(D^0\Psi_{\mathrm{KG0}}(\boldsymbol{x}))$. To write this in the form of an inner product given by equation (35) above, consider (the position representative of) a new vector $\boldsymbol{\lambda}(\boldsymbol{x}) = (\Psi_{\mathrm{KG0}}(\boldsymbol{x}), (m^{-1}\mathrm{i}D^0)\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^T$, instead of $\Psi_{\mathrm{KG0}}(\boldsymbol{x})$. Then

$$j^0_{\mathrm{KG0}} = \boldsymbol{\lambda}(\boldsymbol{x})^\dagger \boldsymbol{K}\boldsymbol{\lambda}(\boldsymbol{x}) \qquad \text{with } \boldsymbol{K} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{44}$$

Hence $\langle\boldsymbol{\lambda}|\boldsymbol{\lambda}\rangle = \int j^0_{\mathrm{KG0}}\,\mathrm{d}^3\boldsymbol{x} = \int \boldsymbol{\lambda}(\boldsymbol{x})^\dagger \boldsymbol{K}\boldsymbol{\lambda}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}$ as required. Observe that j^0_{KG0} is indefinite [10]. What sort of equation does $\boldsymbol{\lambda}(\boldsymbol{x})$ satisfy? Mathematically, the answer is simple. Given a second-order ordinary differential equation for some quantity X, it can be written as two first-order equations, for X and $\dot{X}$ [13]. Hence there appears a two-component equation for the new wavefunction $\boldsymbol{\lambda}(\boldsymbol{x}) = (\Psi_{\mathrm{KG0}}(\boldsymbol{x}), (m^{-1}\mathrm{i}D^0)\Psi_{\mathrm{KG0}}(\boldsymbol{x}))^T$. This is a promising candidate for a relativistic wave equation for spin-0 particles, but from the theory of indefinite inner product spaces [10] it is known that the most convenient form of $\boldsymbol{K}$ is a diagonal matrix, and an indefinite (also Hermitian and non-singular) $\boldsymbol{K}$ can always be rotated to diagonal form via a conjunctive transformation. Hence define $\boldsymbol{\lambda}(\boldsymbol{x}) = S\boldsymbol{\Theta}(\boldsymbol{x})$, with

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{45}$$

A new inner product is obtained, with matrix

$$\boldsymbol{K}' = S^\dagger \boldsymbol{K} S \qquad \boldsymbol{K}' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\boldsymbol{\Theta}(\boldsymbol{x}) = \frac{1}{\sqrt{2}}\begin{pmatrix} (1 + (m^{-1}\mathrm{i}D^0))\Psi_{\mathrm{KG0}}(\boldsymbol{x}) \\ (1 - (m^{-1}\mathrm{i}D^0))\Psi_{\mathrm{KG0}}(\boldsymbol{x}) \end{pmatrix}. \tag{46}$$

$\boldsymbol{\Theta}(\boldsymbol{x})$ is precisely $\Psi_{\mathrm{FV0}}(\boldsymbol{x})$. Hence the natural mathematical construction of an inner product space based upon the conserved current of the KG0 equation leads to the FV0 equation. It is evident that this construction is just the Feshbach–Villars linearization procedure.

The Feshbach–Villars linearization procedure has been justified for spin-0 particles. In the spin-$\frac{1}{2}$ case, one must start with the KG$\frac{1}{2}$ equation. The conserved current density for the KG$\frac{1}{2}$ equation is derived in a similar manner to that for the KG0 equation (equation (40)), with an additional consideration to ensure that j^μ transforms correctly.

Consider a manifestly covariant wave equation of the form $F\psi(\boldsymbol{x}) = 0$, where F contains a mass term multiplied by the identity matrix and also derivative terms multiplied by matrices. Note that equations (15)–(17) are of this form. An equation of the form $\partial_\mu j^\mu = 0$ is sought, and $\partial_\mu j^\mu$ is a scalar. An inner product formalism requires $\langle\psi|\psi\rangle = \int j^0\,\mathrm{d}^3\boldsymbol{x} = \int \psi^\dagger(\boldsymbol{x})\boldsymbol{K}\psi(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}$, i.e. j^0 being the product of some vector with the complex

conjugate transpose of the same vector and a non-singular Hermitian matrix sandwiched in the middle. Hence it is natural to form the quantities $\psi^\dagger(\boldsymbol{x})$ multiplied by $F\psi(\boldsymbol{x})$, and $(F\psi(\boldsymbol{x}))^\dagger$ multiplied by $\psi(\boldsymbol{x})$. In order to obtain products which transform as scalars, in the spin-$\frac{1}{2}$ case a matrix Λ is inserted to give $\psi^\dagger(\boldsymbol{x})\Lambda F\psi(\boldsymbol{x})$ and $(F\psi(\boldsymbol{x}))^\dagger\Lambda\psi(\boldsymbol{x})$. These will transform as scalars if $\psi^\dagger(\boldsymbol{x})\Lambda\psi(\boldsymbol{x})$ transforms as a scalar. The sum or difference of these products is then taken (which removes the mass terms) in order to form the divergence of a 4-vector. One must check that j^μ indeed transforms as a 4-vector, which is trivial if one knows how $\psi(\boldsymbol{x})$ transforms. The KG0 conserved current density given by equation (41) transforms as a 4-vector, as $\Psi_{\mathrm{KG0}}(\boldsymbol{x})$ is a scalar. Here $\Lambda = 1$. The Dirac equation has $F = \mathrm{i}\not{D} - m\mathbf{1}_4$, and $\Lambda = \gamma^0$. The quantity $\Psi_{\mathrm{D}}^\dagger(\boldsymbol{x})\Lambda F\Psi_{\mathrm{D}}(\boldsymbol{x})$ is simply $\Psi_{\mathrm{D}}^\dagger(\boldsymbol{x})(\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4 - H_{\mathrm{D}})\Psi_{\mathrm{D}}(\boldsymbol{x})$, with the Dirac Hamiltonian H_{D} given in [12]. Thus, the relativistic procedure discussed here reduces for the Dirac equation to the same method (equation (38)) used to derive the conserved current density for the Schrödinger equation.

In the KG$\frac{1}{2}$ case Λ is again γ^0, as $\Psi_{\mathrm{KG1/2}}(\boldsymbol{x})$ is also a bispinor. Using equation (32) for the KG$\frac{1}{2}$ equation, construct the quantity

$$\Psi_{\mathrm{KG1/2}}^\dagger(\boldsymbol{x})\gamma^0((D^\mu D_\mu + m^2)\mathbf{1}_4 + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu})\Psi_{\mathrm{KG1/2}}(\boldsymbol{x}) - \left(\left((D^\mu D_\mu + m^2)\mathbf{1}_4 + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\Psi_{\mathrm{KG1/2}}(\boldsymbol{x})\right)^\dagger \gamma^0\Psi_{\mathrm{KG1/2}}(\boldsymbol{x}) = 0. \tag{47}$$

This can be written as $\partial_\mu j^\mu = 0$, with

$$\begin{aligned} j^\mu_{\mathrm{KG1/2}} &= m^{-1}\mathrm{i}(\Psi^\dagger_{\mathrm{KG1/2}}(\boldsymbol{x})\gamma^0 \overleftrightarrow{D^\mu} \Psi_{\mathrm{KG1/2}}(\boldsymbol{x})) \\ &= m^{-1}\mathrm{i}(\Psi^\dagger_{\mathrm{KG1/2}}(\boldsymbol{x})\gamma^0 D^\mu\Psi_{\mathrm{KG1/2}}(\boldsymbol{x}) - (D^\mu\Psi_{\mathrm{KG1/2}}(\boldsymbol{x}))^\dagger\gamma^0\Psi_{\mathrm{KG1/2}}(\boldsymbol{x})). \end{aligned} \tag{48}$$

The definition $\Psi_{\mathrm{KG0}} = \psi\Psi_0 = \psi\otimes\Psi_0$ given in section 5 can now be used to conveniently define the FV$\frac{1}{2}$ inner product. With this definition $j^0_{\mathrm{KG1/2}}$ becomes

$$j^0_{\mathrm{KG1/2}} = m^{-1}\mathrm{i}(\psi^*(\boldsymbol{x}) \overleftrightarrow{D^0} \psi(\boldsymbol{x}))(\overline{\Psi}_0(\boldsymbol{x})\Psi_0(\boldsymbol{x})). \tag{49}$$

To obtain an inner product in the standard form (35) the Feshbach–Villars linearization procedure is applied to the KG$\frac{1}{2}$ equation. Construct

$$\boldsymbol{\lambda}_{1/2}(\boldsymbol{x}) = \begin{pmatrix} \Psi_{\mathrm{KG1/2}}(\boldsymbol{x}) \\ (m^{-1}\mathrm{i}D^0)\Psi_{\mathrm{KG1/2}}(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} \psi(\boldsymbol{x}) \\ (m^{-1}\mathrm{i}D^0)\psi(\boldsymbol{x}) \end{pmatrix} \otimes \Psi_0(\boldsymbol{x}) \tag{50}$$

and perform the same conjunctive transformation as in equation (46) to obtain

$$\boldsymbol{\lambda}_{1/2}(\boldsymbol{x}) = S_{1/2}\boldsymbol{\Theta}_{1/2}(\boldsymbol{x}) = (S\otimes\mathbf{1}_4)\boldsymbol{\Theta}_{1/2}(\boldsymbol{x}) \tag{51}$$

with S given by equation (45) as before. The inner product for $\boldsymbol{\lambda}_{1/2}(\boldsymbol{x})$

$$\langle\boldsymbol{\lambda}_{1/2}|\boldsymbol{\lambda}_{1/2}\rangle = \int \boldsymbol{\lambda}^\dagger_{1/2}(\boldsymbol{x})\boldsymbol{K}_{1/2}\boldsymbol{\lambda}_{1/2}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x} = \int (\boldsymbol{\lambda}^\dagger(\boldsymbol{x})\boldsymbol{K}\boldsymbol{\lambda}(\boldsymbol{x}))(\overline{\Psi}_0(\boldsymbol{x})\Psi_0(\boldsymbol{x}))\,\mathrm{d}^3\boldsymbol{x} \tag{52}$$

is given by

$$\boldsymbol{K}_{1/2} = \boldsymbol{K}\otimes\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\otimes\gamma^0 \tag{53}$$

while for $\boldsymbol{\Theta}_{1/2}(\boldsymbol{x})$

$$\boldsymbol{K}'_{1/2} = \boldsymbol{K}' \otimes \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \gamma^0. \tag{54}$$

$\boldsymbol{\Theta}_{1/2}(\boldsymbol{x})$ is precisely the FV$\frac{1}{2}$ wavefunction $\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x}) = \Psi \otimes \Psi_0$ and

$$\begin{aligned}\langle \boldsymbol{\Theta}_{1/2} | \boldsymbol{\Theta}_{1/2} \rangle = \langle \Psi'_{\mathrm{FV}1/2} | \Psi'_{\mathrm{FV}1/2} \rangle &= \int \Psi'^{\dagger}_{\mathrm{FV}1/2}(\boldsymbol{x}) \boldsymbol{K}'_{1/2} \Psi'_{\mathrm{FV}1/2}(\boldsymbol{x})\, \mathrm{d}^3\boldsymbol{x} \\ &= \int (\Psi^{\dagger}(\boldsymbol{x}) \boldsymbol{K}' \Psi(\boldsymbol{x}))(\overline{\Psi}_0(\boldsymbol{x}) \Psi_0(\boldsymbol{x}))\, \mathrm{d}^3\boldsymbol{x}\end{aligned} \tag{55}$$

defines the inner product for the FV$\frac{1}{2}$ equation.

The Feshbach–Villars linearization procedure has been shown to be a natural mathematical procedure which can be applied to the KG$\frac{1}{2}$ equation as well as the KG0 equation.

Equations (49)–(51) clearly show that the procedure preserves the mathematically convenient bispinor in the wavefunction.

It is evident from the above discussion that the FV$\frac{1}{2}$ equation has a useful advantage over the KG$\frac{1}{2}$ equation in that the FV$\frac{1}{2}$ inner product assumes the standard mathematical form. This, combined with the fact that only the FV$\frac{1}{2}$ equation is in Hamiltonian form, indicates that the FV$\frac{1}{2}$ equation is the preferred equation when the contruction of a relativistic quantum mechanics is considered.

7. The eight-component equation decoupled

In the Weyl representation of the gamma matrices the FV$\frac{1}{2}$ equation decouples into two separate equations, as mentioned above. This can be seen, as discussed in section 3, due to the possibility of forming a second-order relativistically invariant equation using only one of the two spinor irreducible representations. Alternatively, it can be understood by considering the matrix γ_5.

Consider equations (16) and (17) given in section 4. For convenience, define the matrix $\boldsymbol{A} = \mathrm{i}\,\not{\partial}/m$, write equation (16) as $(\boldsymbol{A}^2 - \mathbf{1}_4^2)\Psi = 0$ and equations (17) as $(\boldsymbol{A} \pm \mathbf{1}_4)\Psi_{\pm} = 0$. Let $\mathcal{D}_3$ and $\mathcal{D}_4$ be the subspaces of solutions of the KG$\frac{1}{2}$ equation defined by

$$\Psi_3 \in \mathcal{D}_3,\ \ \Psi_3 = \tfrac{1}{2}(\mathbf{1}_4 + \gamma_5)\Psi \qquad \Psi_4 \in \mathcal{D}_4,\ \ \Psi_4 = \tfrac{1}{2}(\mathbf{1}_4 - \gamma_5)\Psi. \tag{56}$$

Thus Ψ_3 and Ψ_4 are linear combinations of KG$\frac{1}{2}$ solutions. $\mathcal{D}_3$ and $\mathcal{D}_4$ together form a decomposition of $\mathcal{F}$. Ψ_3 and Ψ_4 are eigenstates of $\boldsymbol{A}^2$ with eigenvalue +1 (solutions of equation (16)) . All elements of $\mathcal{D}_3$ (respectively $\mathcal{D}_4$) are eigenvectors of γ_5 with eigenvalue +1 (respectively –1). Hence γ_5 decomposes $\mathcal{F}$ into $\mathcal{D}_3$ and $\mathcal{D}_4$. The existence of invariant subspaces of γ_5 requires a basis in which γ_5 and the projectors $\frac{1}{2}(\mathbf{1}_4 \pm \gamma_5)$ are block diagonal. In this basis the vector Ψ_3 (respectively Ψ_4) has the lower (respectively upper) two components zero. Since both Ψ_3 and Ψ_4 are eigenvectors of $\boldsymbol{A}^2$, then in this basis $\boldsymbol{A}^2$ must also be block diagonal and hence the KG$\frac{1}{2}$ equation (and thus the FV$\frac{1}{2}$ equation) decouples into two parts. Note that elements of $\mathcal{D}_-$ and $\mathcal{D}_+$ are not eigenvectors of γ_5, so a similar result does not follow for the Dirac equation.

The decoupling raises the possibility of using only one of the decoupled equations for calculations. In fact this was done already for hydrogenic atoms in [1]. There, however, the calculation was only carried far enough to obtain the four-component wavefunction Ψ_ξ of one of the decoupled equations. In section 10, the eight-component wavefunctions $\Psi_{\mathrm{FV}1/2}$ are obtained for the part of the solution space describing the bound electron states. In the rest of this section, a general procedure to obtain the full solution space for $\Psi_{\mathrm{FV}1/2}$ will be outlined.

The decoupling of the $\mathrm{FV}\frac{1}{2}$ equation occurs when it is written as in equation (5). It does not matter which of the decoupled equations is used for calculations, but we shall always choose the upper equation ($(\mathrm{i}\frac{\partial}{\partial t}\mathbf{1}_4 - H_\xi)\Psi_\xi = 0$) for consistency. Given a solution Ψ_ξ of the upper half of equation (5), the solution $\Psi_{\dot\eta}$ of the lower half is simple to obtain via spatial inversion. Then it is necessary to combine the two solutions with an appropriate factor $\mathrm{e}^{\mathrm{i}\delta}$ to form the eight-component wavefunction $\Psi_{\mathrm{FV}1/2} = (\Psi_\xi, \mathrm{e}^{\mathrm{i}\delta}\Psi_{\dot\eta})^T$. The point of the factor is the following. The operator γ_5 commutes with the $\mathrm{KG}\frac{1}{2}$ equation. This corresponds to the commutativity of $\mathbf{1}_2 \otimes \gamma_5$ with $H'_{\mathrm{FV}1/2}$ (equation (30)). Specializing to the Weyl representation of the gamma matrices where γ_5 is given by equation (11), and writing the $\mathrm{FV}\frac{1}{2}$ equation as equation (5) to take advantage of the decoupling, leads to the result that if $(\Psi_\xi, \Psi_{\dot\eta})^T$ is a solution of equation (5), then so is $(\Psi_\xi, -\Psi_{\dot\eta})^T$. Both these solutions have the same eigenvalue E. Since a linear combination is also a solution with eigenvalue E, then, in general, $(\Psi_\xi, \mathrm{e}^{\mathrm{i}\delta}\Psi_{\dot\eta})^T$ is a solution of equation (5), also with eigenvalue E. In describing the solution space a linearly independent set of solutions is desired and thus two and only two choices of δ are necessary.

To decide which two values of δ should be used, one only has to consider how $\Psi_{\dot\eta}$ is obtained from Ψ_ξ. All the quantities in the wavefunction Ψ_ξ are simply replaced by their corresponding spatially inverted counterparts. Clearly, the same must be done in going from $\Psi_{\dot\eta}$ to Ψ_ξ. The process $\Psi_\xi \to \Psi_{\dot\eta} \to \Psi_\xi$ is the identity and so $\mathrm{e}^{2\mathrm{i}\delta} = +1$. Therefore, $\delta = 0$ and $\delta = \pi$. In an arbitrary gamma matrix representation the $\delta = \pi$ solution is obtained from the $\delta = 0$ solution by the premultiplication of Ψ_0 by γ_5.

After the two solutions are obtained they have to be normalized. The value of the normalization integral of the $\delta = \pi$ solution is always opposite in sign to that of the $\delta = 0$ solution. As discussed in the next section, for a given solution, the normalization process itself cannot change the sign of the normalization integral. For square-integrable states, a solution with $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle > 0$ can be 'maximally' (see next section) normalized to give $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle = +1$, while a solution with $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle < 0$ can be 'maximally' normalized to give $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle = -1$. In the case of continuum states where the wavefunction is not square-integrable, the normalization takes the form $\langle\Psi_{\mathrm{FV}1/2\ E}|\Psi_{\mathrm{FV}1/2E'}\rangle = +\delta(E - E')$ for a $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle > 0$ solution and $\langle\Psi_{\mathrm{FV}1/2\ E}|\Psi_{\mathrm{FV}1/2E'}\rangle = -\delta(E - E')$ for a $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle < 0$ solution.

It is in general not true that the $\delta = 0$ states are the ones with $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle > 0$. However, an eigenvector $\Psi_{\mathrm{FV}1/2}$ of $H_{\mathrm{FV}1/2}$ with eigenvalue $E > 0$ and $\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle > 0$ is a state with charge e and positive energy which describes the usual 'particle' state and reduces in the non-relativistic limit to the usual Schrödinger state. The charge e is seen in the minimal coupling $\partial_\mu \to \partial_\mu + \mathrm{i}eA_\mu$.

After the eight-component solutions are obtained and maximally normalized, their physical interpretation can be discussed. The procedure outlined in this section for constructing $\Psi_{\mathrm{FV}1/2}$ from Ψ_ξ clearly holds for any problem in which a classical external electromagnetic field is introduced via a minimal coupling.

8. The associated quantum mechanical formalism for the FV$\frac{1}{2}$ equation

Feshbach and Villars [3] give some results to develop a quantum mechanical formalism for the FV0 equation using the indefinite inner product $\langle\Psi_{\mathrm{FV0}}|\Psi_{\mathrm{FV0}}\rangle = \int \Psi^{\dagger}_{\mathrm{FV0}}(\boldsymbol{x})\tau_3\Psi_{\mathrm{FV0}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}$. The presence of the matrix $\boldsymbol{K}' = \tau_3$ leads to a modified definition of $\Omega^{\sharp}$ in section 6, namely (using $\tau_3 = \tau_3^{\dagger} = \tau_3^{-1}$) $\Omega^{\sharp} = \tau_3\Omega^{\dagger}\tau_3$, whereas for the unitary inner product it is given by $\Omega^{\sharp} = \Omega^{\dagger}$. In general, the matrix $\boldsymbol{K}$ in (35) must be non-singular and Hermitian [10] and this is the case for both $\boldsymbol{K}' = \tau_3$ and $\boldsymbol{K}'_{1/2} = \tau_4$ in (54). Consider now the spin-$\frac{1}{2}$ case, where $\tau_4 = \tau_4^{\dagger} = \tau_4^{-1}$. A operator is pseudo-Hermitian if $\tau_4\Omega^{\dagger}\tau_4 = \Omega$ and a transformation U is pseudo-unitary if $\tau_4U^{\dagger}\tau_4 = U^{-1}$. $H'_{\mathrm{FV1/2}}$ is easily shown to be pseudo-Hermitian by using properties of the Kronecker product $(A\otimes B)^{\dagger} = A^{\dagger}\otimes B^{\dagger}$ and $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$ which are valid here as the elements of the matrix τ_4 are c-numbers, and the dimensions of the matrices appropriate.

$$\tau_4H'^{\dagger}_{\mathrm{FV1/2}}\tau_4 = (\tau_3(\tau_3+\mathrm{i}\tau_2)^{\dagger}\tau_3)\otimes\left(\gamma^0\left(\frac{1}{2m}\left(-\boldsymbol{D}^2\mathbf{1}_4+\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\right)^{\dagger}\gamma^0\right) + m^*(\tau_3\tau_3^{\dagger}\tau_3)\otimes(\gamma^0(\mathbf{1}_4)^{\dagger}\gamma^0) + (eA_0)^*(\tau_3(\mathbf{1}_2)^{\dagger}\tau_3)\otimes(\gamma^0(\mathbf{1}_4)^{\dagger}\gamma^0) \tag{57}$$

which equals $H'_{\mathrm{FV1/2}}$, as $\sigma^{\mu\nu\dagger} = \gamma^0\sigma^{\mu\nu}\gamma^0$ and m, e and A_0 are real. Hence $H'_{\mathrm{FV1/2}}$ (analogously to H_{FV0}) plays the rôle of the time transformation operator. $H'_{\mathrm{FV1/2}}$ has a complete set of orthogonal eigenfunctions, as will be shown in a later paper. These are obtained simply by solving the equation $H_\xi\Psi_\xi = E\Psi_\xi$ and then constructing the full FV$\frac{1}{2}$ wavefunctions using the method given in the previous section with $\delta = 0$ and $\delta = \pi$ respectively. Thus a quantum mechanical formalism similar to that for the FV0 equation [3] can be developed, for example $\langle\mathrm{d}\Omega/\mathrm{d}t\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\Omega\rangle$. The velocity operator for free particles is given by $v_k = (p_k/m)((\tau_3+\mathrm{i}\tau_2)\otimes\mathbf{1}_4)$ which is closely related to the velocity operator of the FV0 equation $v_k = (p_k/m)(\tau_3+\mathrm{i}\tau_2)$ [3].

In non-relativistic quantum mechanics the expectation value of an operator Ω is defined by $\langle\Omega\rangle = (\int\Psi^*(\boldsymbol{x})\Omega(\boldsymbol{x})\Psi(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x})/(\int\Psi^*(\boldsymbol{x})\Psi(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x})$. The definition given in [3] for the FV0 equation

$$\langle\Omega\rangle = \int\Psi^{\dagger}_{\mathrm{FV0}}(\boldsymbol{x})\tau_3\Omega(\boldsymbol{x})\Psi_{\mathrm{FV0}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x} \tag{58}$$

clearly holds only for normalized wavefunctions and implies, given the indefinite inner product,

$$\langle\Omega\rangle = \frac{\int\Psi^{\dagger}_{\mathrm{FV0}}(\boldsymbol{x})\tau_3\Omega(\boldsymbol{x})\Psi_{\mathrm{FV0}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}}{|\int\Psi^{\dagger}_{\mathrm{FV0}}(\boldsymbol{x})\tau_3\Psi_{\mathrm{FV0}}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}|} \tag{59}$$

in the general case. Equation (59) must also be used for plane waves and other solutions which are not square-integrable. While the definition (58) is always used in discussions of the FV0 equation both in [3] and in textbooks (see for example [14]), and a consistent interpretation requires this, some mathematical justification from the theory of indefinite inner product spaces is helpful. This is found in the normalization of vectors that have $\langle\psi|\psi\rangle < 0$ (see chapter IX of [10]). There it is shown that the most that can be done is to 'maximally normalize' such vectors, which means dividing, not by $\langle\psi|\psi\rangle$ as in the case

of positive definite inner products, but by $|\langle\psi|\psi\rangle|$. Thus follows the definition given by equation (59). For the FV$\frac{1}{2}$ equation it becomes

$$\langle\Omega\rangle_{\mathrm{FV}1/2} = \frac{\int {\Psi'_{\mathrm{FV}1/2}}^{\dagger}(\boldsymbol{x})\tau_4\Omega(\boldsymbol{x})\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}}{\left|\int {\Psi'_{\mathrm{FV}1/2}}^{\dagger}(\boldsymbol{x})\tau_4\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x}\right|}. \tag{60}$$

From the FV$\frac{1}{2}$ Hamiltonian (30) it is clear that the spin term $\sigma^{\mu\nu}F_{\mu\nu}$ in (32) is not broken by the Feshbach–Villars linearization procedure in an arbitrary gamma matrix representation. Corresponding to this the bispinor is preserved in the wavefunction. The gamma matrix algebra together with its unitary transformations among representations is thus retained. In general, however, a transformation which preserves the value of $\langle\Psi'_{\mathrm{FV}1/2}|\Psi'_{\mathrm{FV}1/2}\rangle$ for any $|\Psi'_{\mathrm{FV}1/2}\rangle$ is pseudo-unitary with $\tau_4U^{\dagger}\tau_4 = U^{-1}$. The matrix $\tau_4 = \tau_3 \otimes \gamma^0$ and it is the τ_3 which makes the transformation pseudo-unitary analogous to the spin-0 case. If we consider a change of the gamma matrix representation according to the usual unitary transformations, this acts only on the bispinor part Ψ_0 of $\Psi_{\mathrm{FV}1/2} = \Psi \otimes \Psi_0$, and clearly also preserves the inner product with $j^0_{\mathrm{FV}1/2}(\boldsymbol{x}) = {\Psi'_{\mathrm{FV}1/2}}^{\dagger}(\boldsymbol{x})\tau_4\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x}) = (\Psi^{\dagger}(\boldsymbol{x})\tau_3\Psi(\boldsymbol{x}))(\overline{\Psi_0}(\boldsymbol{x})\Psi_0(\boldsymbol{x}))$, because bilinear covariants are invariant under such transformations. In section 6 it was mentioned that the ideal form of an indefinite inner product is a diagonal matrix (with 1s and –1s on the diagonal). $\tau_4 = \tau_3 \otimes \gamma^0$ is block diagonal. However, the γ^0 is clearly understood to be due to the gamma matrix algebra present in any correct relativistic spin-$\frac{1}{2}$ equation, and the τ_3 part of τ_4 is of the desired diagonal form diag(+1, –1). The identity operator for the FV$\frac{1}{2}$ equation is $\sum_i \sigma_i|\Psi'_{\mathrm{FV}1/2i}\rangle\langle\Psi'_{\mathrm{FV}1/2i}|$, where $\sigma_i = \pm 1$ if $\langle\Psi'_{\mathrm{FV}1/2i}|\Psi'_{\mathrm{FV}1/2i}\rangle = \pm 1$.

9. The use of Kronecker products for the FV$\frac{1}{2}$ equation

There are two distinct yet equivalent ways of writing the FV$\frac{1}{2}$ equation using Kronecker products. The first takes advantage of the decoupling of the equation in the Weyl representation of the gamma matrices, while the second preserves the bispinor in the FV$\frac{1}{2}$ wavefunction. Consider the first case where the FV$\frac{1}{2}$ wavefunction is

$$\Psi_{\mathrm{FV}1/2} = \begin{pmatrix} \Psi_{\xi} \\ \Psi_{\dot{\eta}} \end{pmatrix} = \begin{pmatrix} \Psi \otimes \xi_0 \\ \Psi \otimes \dot{\eta_0} \end{pmatrix}. \tag{61}$$

This was used in equations (23) to (27) to define the linearization procedure for the two decoupled parts of the KG$\frac{1}{2}$ equation in the Weyl representation of the gamma matrices. The FV$\frac{1}{2}$ inner product defined via equation (54) is valid for the second case with $\Psi_{\mathrm{FV}1/2} = \Psi \otimes \Psi_0$, rather than for equation (61). If equation (61) is used, the inner product becomes

$$\langle\Psi_{\mathrm{FV}1/2}|\Psi_{\mathrm{FV}1/2}\rangle = \int \Psi^{\dagger}_{\mathrm{FV}1/2}(\boldsymbol{x})\tau_5\Psi_{\mathrm{FV}1/2}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x} \tag{62}$$

$$\tau_5 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 & \mathbf{0} & \mathbf{0} \end{pmatrix} = \tau_1 \otimes (\tau_3 \otimes \mathbf{1}_2). \tag{63}$$

Note that τ_5 is non-singular and Hermitian as required [10]. With this value of τ_5,

$$\langle \Psi_{\mathrm{FV}1/2} | \Psi_{\mathrm{FV}1/2} \rangle = \int (\Psi^\dagger(\boldsymbol{x})\tau_3\Psi(\boldsymbol{x}))(\xi_0^\dagger(\boldsymbol{x})\dot{\eta}_0(\boldsymbol{x}) + \dot{\eta}_0{}^\dagger(\boldsymbol{x})\xi_0(\boldsymbol{x}))\,\mathrm{d}^3\boldsymbol{x}. \quad (64)$$

Consider now the method to obtain the $\mathrm{FV}\frac{1}{2}$ inner product containing τ_5 from $\int j^0_{\mathrm{KG}1/2}\,\mathrm{d}^3\boldsymbol{x}$, where

$$j^0_{\mathrm{KG}1/2} = m^{-1}\mathrm{i}(\Psi^\dagger_{\mathrm{KG}1/2}(\boldsymbol{x})\gamma^0 \overleftrightarrow{D^0} \Psi_{\mathrm{KG}1/2}(\boldsymbol{x})). \quad (65)$$

To construct $\Psi_{\mathrm{FV}1/2}$ given by equation (61), firstly the bispinor $\Psi_{\mathrm{KG}1/2}$ is split into ξ and $\dot{\eta}$ (see equation (23)), hence γ^0 in equation (65) is split into $\tau_1 \otimes \mathbf{1}_2$ ($\gamma^0 = \tau_1 \otimes \mathbf{1}_2$ in the Weyl representation of the gamma matrices). Then

$$\Psi_{\mathrm{KG}1/2} = \begin{pmatrix} \xi \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} \psi\xi_0 \\ \psi\dot{\eta}_0 \end{pmatrix} = \begin{pmatrix} \psi \otimes \xi_0 \\ \psi \otimes \dot{\eta}_0 \end{pmatrix} \quad (66)$$

is rewritten using the linearization procedure as $\Psi_{\mathrm{FV}1/2} = (\Psi \otimes \xi_0, \Psi \otimes \dot{\eta}_0)^T$. Hence $\tau_1 \otimes \mathbf{1}_2 \rightarrow \tau_1 \otimes (\tau_3 \otimes \mathbf{1}_2) = \tau_5$. In general, the linearization procedure takes a j^0 containing a derivative term for a second-order equation and replaces the derivative by a τ_3 for a first-order (in the time derivative) equation with twice the number of components. In the spin-$\frac{1}{2}$ case, it is necessary to write τ_3 together with γ^0 in a convenient fashion.

If instead the $\mathrm{KG}\frac{1}{2}$ equation is written in an arbitrary gamma matrix representation and the linearization procedure applied directly, then this corresponds to the second case with the $\mathrm{FV}\frac{1}{2}$ wavefunction defined by $\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x}) = \Psi \otimes \Psi_0$ and

$$\langle \Psi'_{\mathrm{FV}1/2} | \Psi'_{\mathrm{FV}1/2} \rangle = \int \Psi'^\dagger_{\mathrm{FV}1/2}(\boldsymbol{x})\tau_4\Psi'_{\mathrm{FV}1/2}(\boldsymbol{x})\,\mathrm{d}^3\boldsymbol{x} \quad (67)$$

where τ_4 is $\boldsymbol{K}'_{1/2}$ defined by equation (54). Specializing to the Weyl representation of the gamma matrices for comparison purposes, this becomes

$$\tau_4 = \begin{pmatrix} \mathbf{0} & \mathbf{1}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{1}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{1}_2 \\ \mathbf{0} & \mathbf{0} & -\mathbf{1}_2 & \mathbf{0} \end{pmatrix} = \tau_3 \otimes \gamma^0_{\mathrm{Weyl}} = \tau_3 \otimes (\tau_1 \otimes \mathbf{1}_2). \quad (68)$$

Here the bispinor is not split so γ^0 remains itself in τ_4. Equation (67) in the Weyl representation of the gamma matrices gives of course exactly the right-hand side of equation (64).

Consider what happens if an attempt is made to derive τ_5 in an arbitrary gamma matrix representation. Here, τ_4 is $\tau_3 \otimes \gamma^0$. If it is possible that γ^0 can be written as $\gamma^0 = \tau_6 \otimes \tau_7$, where τ_6 and τ_7 reflect the splitting of the bispinor $\Psi_{\mathrm{KG}1/2}$ into $(\alpha, \beta)^T$ (equation (29)), then one could derive $\tau_5 = \tau_6 \otimes (\tau_3 \otimes \tau_7)$. Only in the Weyl representation of the gamma matrices is it convenient to do this. Both methods obviously give the same final $\langle \Psi | \Psi \rangle$ as both inner products are derived from the same $j^0_{\mathrm{KG}1/2}$. To obtain the wavefunctions in the two cases from each other, one only needs to know ξ_0, $\dot{\eta}_0$ and Ψ. In general, the quantum mechanical formalism can be developed for either of the two ways of writing the $\mathrm{FV}\frac{1}{2}$ equation using Kronecker products. They are equivalent to one another. Convenience dictates which is chosen in a given situation.

While the use of Kronecker products is convenient, some care must be taken when operators are involved. For example, equation (30) for the FV$\frac{1}{2}$ Hamiltonian $H'_{\mathrm{FV}1/2}$ in an arbitrary gamma matrix representation contains the term $(\tau_3 + \mathrm{i}\tau_2) \otimes (-\boldsymbol{D}^2\mathbf{1}_4/(2m))$. When acting on the wavefunction $\Psi'_{\mathrm{FV}1/2} = \Psi \otimes \Psi_0$, the operator $\boldsymbol{D}^2$ of course acts on the $\boldsymbol{x}$ dependent parts in the whole wavefunction, not just those in Ψ_0. Strictly speaking, the term should be written as $(-\boldsymbol{D}^2/(2m))((\tau_3 + \mathrm{i}\tau_2) \otimes \mathbf{1}_4)$, but this is messy when added to the term $(\tau_3 + \mathrm{i}\tau_2) \otimes (e\sigma^{\mu\nu}F_{\mu\nu}/(4m))$ which itself is correct (but could be written as $(eF_{\mu\nu}/(4m))((\tau_3 + \mathrm{i}\tau_2) \otimes \sigma^{\mu\nu})$). In all cases in this series of papers, the use of Kronecker products to write a 8×8 matrix $\boldsymbol{M}$ as $\boldsymbol{A} \otimes \boldsymbol{B}$ where $\boldsymbol{A}$ and $\boldsymbol{B}$ are 2×2 and 4×4 matrices respectively is rigorously correct, as are the definitions of the wavefunctions Ψ in (24) and $\boldsymbol{\lambda}_{1/2}(\boldsymbol{x})$ in (50) (there is no time dependence in Ψ_0). For convenience, however, sometimes a scalar (in the eight-dimensional space) operator s will be placed in $\boldsymbol{M}$ as $\boldsymbol{A} \otimes (s\boldsymbol{B})$ or $(s\boldsymbol{A}) \otimes \boldsymbol{B}$ or $\boldsymbol{A} \otimes (\boldsymbol{B}s)$ or $(\boldsymbol{A}s) \otimes \boldsymbol{B}$, instead of $s(\boldsymbol{A} \otimes \boldsymbol{B})$. These comments also apply to the Weyl representation of the gamma matrices, where a 4×4 matrix is written as the Kronecker product of 2×2 matrices.

10. The hydrogen atom

The solution to the hydrogen atom was discussed at length in the first paper to suggest that the FV$\frac{1}{2}$ equation can be used successfully for problems of a single particle moving in a classical external electromagnetic field. It was seen there that only one of the two equations is necessary to obtain the eigenvalue spectrum and four-component wavefunctions Ψ_ξ. The full FV$\frac{1}{2}$ wavefunction $\Psi_{\mathrm{FV}1/2}$ requires both Ψ_ξ and $\Psi_{\dot\eta}$, together with a choice of the factor $\mathrm{e}^{\mathrm{i}\delta}$. Once $\Psi_{\mathrm{FV}1/2}$ is constructed, then the quantum mechanical formalism with the indefinite inner product can be used to analyse the solutions.

Since $\Psi_{\dot\eta}$ is obtained from Ψ_ξ by spatial inversion, the radial wavefunctions for $\Psi_{\dot\eta}$ must be identical to those of Ψ_ξ and the angular wavefunctions must change by a sign of κ, with κ defined by equation (21), [1]. The difference in κ can be traced back to the angular operator Ω in the preceding paper where $\Omega_\xi = \boldsymbol{L}^2\mathbf{1}_2 - \mathrm{i}Z\alpha\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}}$ becomes $\Omega_{\dot\eta} = \boldsymbol{L}^2\mathbf{1}_2 + \mathrm{i}Z\alpha\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}}$ due to the sign change of the $\boldsymbol{\sigma}\cdot\boldsymbol{E}$ term in the Hamiltonian. Ω_ξ has the same eigenvalues as $\Omega_{\dot\eta}$ leading to the same radial equations and hence energy spectrum.

To normalize the wavefunctions one can use either $\Psi_{\mathrm{FV}1/2} = (\Psi_\xi, \Psi_{\dot\eta})^T$ or $\Psi'_{\mathrm{FV}1/2} = \Psi \otimes \Psi_0$, as described in the previous section. For the hydrogen atom one has

$$\Psi = \begin{pmatrix} \phi(r) \\ \chi(r) \end{pmatrix} \qquad \Psi_0 = \begin{pmatrix} \xi_0 \\ \mathrm{e}^{\mathrm{i}\delta}\dot\eta_0 \end{pmatrix} = \begin{pmatrix} \Theta_{(l,l',j,m)\xi}(\theta,\phi) \\ \mathrm{e}^{\mathrm{i}\delta}\Theta_{(l,l',j,m)\dot\eta}(\theta,\phi) \end{pmatrix} = \begin{pmatrix} \Theta_\xi \\ \mathrm{e}^{\mathrm{i}\delta}\Theta_{\dot\eta} \end{pmatrix}. \tag{69}$$

In the bound-state case, choosing $\delta = 0$ leads to $\langle\Psi'_{\mathrm{FV}1/2}|\Psi'_{\mathrm{FV}1/2}\rangle > 0$, where $\langle\Psi'_{\mathrm{FV}1/2}|\Psi'_{\mathrm{FV}1/2}\rangle$ equals

$$\int (\Psi^\dagger\tau_3\Psi)(\overline{\Psi}_0\Psi_0)\,\mathrm{d}^3\boldsymbol{x} = \int r^2\,\mathrm{d}r(|\phi(r)|^2 - |\chi(r)|^2)\iint \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi(\Theta_\xi^\dagger\Theta_{\dot\eta} + \Theta_{\dot\eta}^\dagger\Theta_\xi). \tag{70}$$

Since $\langle\Psi'_{\mathrm{FV}1/2}|\Psi'_{\mathrm{FV}1/2}\rangle > 0$ and the bound-state solutions are normalizable, the integral in equation (70) is set equal to $+1$, which provides the values of the normalization constants C and C' mentioned in equation (52) of the preceding paper. It is of course possible to define only one normalization constant, but for convenience C' is defined so that the angular

integral equals $+1$, and then C used to normalize the radial integral to $+1$. The values of C and C' are

$$C = \frac{|\Lambda|\sqrt{\Gamma(2\gamma + n' + 1)}}{\Gamma(2\gamma + 2)\sqrt{Z(n'-1)!}} \qquad C' = (2(1-\kappa^2))^{-\frac{1}{2}}. \tag{71}$$

For these states, $\langle H'_{\mathrm{FV}1/2}\rangle = +E$ and so they are interpreted as the usual positive energy electrons with charge $(+e)$ bound in the attractive nuclear potential of charge $(-Ze)$. There is a one-to-one correspondence between these solutions and the Dirac bound-state spectrum.

There are also continuum solutions of the $\mathrm{FV}\frac{1}{2}$ equation, which can be obtained in analogy to the method for the Dirac equation presented in [5]. This amounts to replacing $|\Lambda|$ (equation (28), [1]) by $-\mathrm{i}\Lambda$, $\Lambda \in \Re$, which is necessary as now $|E| > \alpha^{-2}$, and reconsidering the normalization of the wavefunctions. When $E > \alpha^{-2}$, the $\delta = 0$ solutions correspond to the positive energy continuum electron solutions of the Dirac equation. When $E < -\alpha^{-2}$, the $\delta = \pi$ solutions correspond to the negative energy continuum electron solutions of the Dirac equation.

The three sets of solutions discussed above (continuum electrons of both positive and negative energy and bound electrons) exhaust the possible solutions of the Dirac equation. However, they cover only one half of the $\mathrm{FV}\frac{1}{2}$ solution space and the other half has yet to be discussed. For each solution of the Dirac equation with eigenvalue E, there are two linearly independent solutions of the $\mathrm{FV}\frac{1}{2}$ equation having the same eigenvalue E. In the above, δ has been chosen to obtain the solution which corresponds physically to the Dirac solution. The other choice of δ and the physical interpretation it leads to will be considered in a later paper.

It was mentioned in the preceding paper that the values of n' and γ differ for the $j = l + \frac{1}{2}$ cases using the two equations. Here, $\gamma_{\mathrm{FV}1/2} = \gamma_{\mathrm{D}} - 1$ and $n'_{\mathrm{FV}1/2} = n'_{\mathrm{D}} + 1$. The sum $\gamma + n'$ is the same for both equations for a given state, as could be expected. Moreover, the physical interpretation of a given state is identical. The deviation is an artifact of the different radial wavefunctions for the two equations, the correspondence of which will now be discussed.

The wavefunctions derived in the preceding paper can be compared with the Dirac wavefunctions [15] which are (using the standard representation of the gamma matrices)

$$\Psi_{\mathrm{D}jm} = \begin{pmatrix} g(r)\mathcal{Y}^m_{lj} \\ \mathrm{i}f(r)\mathcal{Y}^m_{l'j} \end{pmatrix} \tag{72}$$

$$f(r) = \frac{1}{r}\sqrt{1-\epsilon}\,\mathrm{e}^{-\lambda_{\mathrm{D}}r}(\phi_1 - \phi_2) \tag{73}$$

$$g(r) = \frac{1}{r}\sqrt{1+\epsilon}\,\mathrm{e}^{-\lambda_{\mathrm{D}}r}(\phi_1 + \phi_2) \tag{74}$$

$$\phi_1 = -cn'\left(\frac{Z\alpha}{\sqrt{1-\epsilon^2}} - \chi\right)^{-1/2}(2\lambda_{\mathrm{D}}r)^{\gamma}F(1-n', 2\gamma+1; (2\lambda_{\mathrm{D}}r)) \tag{75}$$

$$\phi_2 = c\left(\frac{Z\alpha}{\sqrt{1-\epsilon^2}} - \chi\right)^{+1/2}(2\lambda_{\mathrm{D}}r)^{\gamma}F(-n', 2\gamma+1; (2\lambda_{\mathrm{D}}r)) \tag{76}$$

where $\chi = \pm(j + \frac{1}{2})$ for $l = j \pm \frac{1}{2}$; $\lambda_{\mathrm{D}} = |\Lambda|$ and c is a normalization constant.

The FV$\frac{1}{2}$ wavefunctions were derived in the Weyl representation of the gamma matrices, whereas the Dirac wavefunctions are given in the standard representation. In general, the Dirac wavefunctions for a central field problem will be of the form given on page 54, [12]

$$\Psi_{\mathrm{D}jm} = \begin{pmatrix} g_j^+(r)\mathcal{Y}_{lj}^m + g_j^-(r)\mathcal{Y}_{l'j}^m \\ \mathrm{i}f_j^+(r)\mathcal{Y}_{lj}^m + \mathrm{i}f_j^-(r)\mathcal{Y}_{l'j}^m \end{pmatrix}. \tag{77}$$

The standard representation of the gamma matrices has the bispinor given by

$$\Psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = S\begin{pmatrix} \xi \\ \dot{\eta} \end{pmatrix} \qquad S = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{1}_2 & \mathbf{1}_2 \\ \mathbf{1}_2 & -\mathbf{1}_2 \end{pmatrix}. \tag{78}$$

Under spatial inversion, α and β transform into themselves (multiplied by opposite phases) and hence this requires either g_j^+ or g_j^- to be zero (and correspondingly f_j^- or f_j^+ to be zero), which gives the form of equation (72). In the Weyl representation of the gamma matrices, ξ and $\dot{\eta}$ transform into each other, and so both $\mathcal{Y}_{lj}^m$ and $\mathcal{Y}_{l'j}^m$ must be included in each part of the wavefunction. If the FV$\frac{1}{2}$ wavefunction is transformed to the standard representation, then this involves only the application of $S(=S^{-1})$ to Ψ_0. This gives

$$\Psi_0' = \sqrt{2}C_1\begin{pmatrix} \mathcal{Y}_{lj}^m \\ \mathrm{i}\kappa\mathcal{Y}_{l'j}^m \end{pmatrix} \tag{79}$$

which shows that, in the same gamma matrix representation, the angular eigenfunctions are identical. Thus the well established classification of states in terms of angular momentum is preserved.

The radial wavefunctions differ in the confluent hypergeometric functions used. The FV$\frac{1}{2}$ equation has only one confluent hypergeometric function for a given state, however this is multiplied by a simple polynomial of the form $1 + m^{-1}\mathrm{i}D^0$, with D^0 containing a r^{-1} piece from the A^0 term. On the other hand, the Dirac wavefunction has a linear combination of two confluent hypergeometric functions.

In understanding the FV$\frac{1}{2}$ wavefunctions, it is useful to compare them to those obtained using the KG$\frac{1}{2}$ equation [16]. In [16] one half of the KG$\frac{1}{2}$ equation is used to derive the solutions, the equation for $\dot{\eta}$. A single second-order radial equation is obtained, which is written in the form of a differential equation for the confluent hypergeometric function. The eigenvalue spectrum and wavefunctions are then derived by analogy with the Schrödinger radial equation. The wavefunctions are not normalized, as no inner product or full KG$\frac{1}{2}$ wavefunction is defined.

The energy eigenvalue spectra derived in [16] are identical to those using the FV$\frac{1}{2}$ (and Dirac) equation. It will now be shown that the FV$\frac{1}{2}$ wavefunctions are consistent with those obtained in [16]. To do this, a connection must be made between the notation in [16] and in [1]. The symbols in [16] are given an AB subscript, and they use units where $\overline{h} = 1 = c$, and E and r are in ordinary units. The following relations exist

$$\gamma_{AB} = Z\alpha \qquad \lambda_{AB} = \frac{Z\epsilon}{|\Lambda|} = n' + \gamma \qquad \rho_{AB} = \rho \qquad S_{AB} = \gamma_{\mathrm{spin0}} \tag{80}$$

$$S_{\pm AB} = \gamma_{\mathrm{spin1/2}} \qquad x_{\pm AB} = \frac{\mathrm{i}\kappa_{\dot{\eta}}}{Z\alpha} \qquad (\kappa_{\dot{\eta}} = -\kappa_{\xi}). \tag{81}$$

The solution method in [16] begins with the standard representation of the gamma matrices and then after some algebra the $\dot{\eta}$ solution is obtained. The angular functions in [16] are identical to $\Theta_{\dot{\eta}}$.

Now consider the radial wavefunctions of the $\mathrm{FV}\frac{1}{2}$ equation. They can be written as

$$\frac{1}{r}\begin{pmatrix} f(r) \\ g(r) \end{pmatrix} = C|\Lambda|\,\mathrm{e}^{-|\Lambda|r}(2|\Lambda|r)^{\gamma}F(1-n',2\gamma+2;2|\Lambda|r) \times \begin{pmatrix} 1+\alpha^2(E+Z/r) \\ 1-\alpha^2(E+Z/r) \end{pmatrix}. \tag{82}$$

The factors $1\pm\alpha^2(E+Z/r)$ are simply $1\pm m^{-1}\mathrm{i}D^0$. Hence, given how Ψ is derived from ψ (equation (24)), one then expects the radial wavefunction derived in [16] to be

$$\frac{1}{r}h(r) = \frac{1}{\sqrt{2}}\frac{1}{r}(f(r)+g(r)) = \sqrt{2}C|\Lambda|\,\mathrm{e}^{-|\Lambda|r}(2|\Lambda|r)^{\gamma}F(1-n',2\gamma+2;2|\Lambda|r). \tag{83}$$

The radial wavefunctions given in [16] are

$$V^{\pm}(\rho) = \mathrm{e}^{-\rho/2}\rho^{S_{\pm}}L^{2S_{\pm}+1}_{\lambda-S_{\pm}-1}(\rho). \tag{84}$$

With $\rho = 2|\Lambda|r$, $L^{\alpha}_{\beta}(\rho) = c'F(-\beta,\alpha+1;\rho)$ [17], $\lambda - S_{\pm} = n'$, $2S_{\pm}+1 = 2\gamma+1$ and c' a constant, the wavefunctions are identical (up to the normalization constant which is not given in [16]). Hence the hydrogen atom results in [1] are consistent with those obtained using the $\mathrm{KG}\frac{1}{2}$ equation.

The claim is made [16] that the Dirac wavefunctions can then be obtained by premultiplying by $((2m)^{-1}(\mathrm{i}\not{D}+m\mathbf{1}_4))$. A somewhat tedious calculation using recurrence relations [18] between different confluent hypergeometric functions shows that this is indeed true. Equation (20) is a general relation linking a solution of the $\mathrm{KG}\frac{1}{2}$ equation with a solution of the Dirac equation. In the case of hydrogenic atoms, this projection produces exactly the Dirac wavefunction from the corresponding $\mathrm{KG}\frac{1}{2}$ wavefunction. This, together with the consistency between the $\mathrm{FV}\frac{1}{2}$ and $\mathrm{KG}\frac{1}{2}$ wavefunctions, suggests that the solution method in [1] combined with the procedure in section 7 is correct. It also demonstrates the relationship between the $\mathrm{FV}\frac{1}{2}$ and Dirac bound-state electron wavefunctions.

Equation (18) shows that any solution of the Dirac equation is also a solution of the $\mathrm{KG}\frac{1}{2}$ equation. However, neither the solution method in section 7 nor that in [16] for the $\mathrm{KG}\frac{1}{2}$ equation gives the Dirac wavefunction. This is easily seen by comparing (83) and (72) where the $\mathrm{KG}\frac{1}{2}$ radial wavefunction is common to all four components in (83) while the Dirac radial wavefunction differs between the upper and lower components in (72). Thus, it is incorrect to take the Dirac wavefunction as the $\mathrm{KG}\frac{1}{2}$ wavefunction and rewrite it using the method in section 5 to obtain the $\mathrm{FV}\frac{1}{2}$ wavefunction. Rather, it is $\Psi_{\mathrm{KG}1/2}$ given by $((2m)^{-1}(\mathrm{i}\not{D}+m\mathbf{1}_4))\Psi_{\mathrm{KG}1/2} = \Psi_{\mathrm{D}}$ which provides the link to the $\mathrm{FV}\frac{1}{2}$ wavefunction.

It was mentioned at the beginning of the solution to the hydrogen atom in [1] that the solution method was valid also for spin-0 with only minor modifications. To obtain the spin-0 results, all that is necessary is to replace γ from equation (35), [1] by equation (36), [1], and the radial functions will be analogous to equation (51), [1]. The angular functions are obtained simply by replacing $\Theta_{(l,l',j,m)}(\theta,\phi)$ (equation (20), [1]) by $Y_{lm}(\theta,\phi)$. Finally, the energy spectrum is the same as that given by equation (44), [1], with γ from equation (36), [1]. The possibility to obtain the exact spin-0 results directly from the exact spin-$\frac{1}{2}$ results is due to the similarity of $H'_{\mathrm{FV}1/2}$ and $H_{\mathrm{FV}0}$. Such a comparison is not available using the Dirac equation. The relationship between the energy spectra of the KG0 equation and the Dirac equation can only be given approximately in an expansion of terms of increasing powers of

$1/m$ (see for example, [12]). Given the similarity of the FV$\frac{1}{2}$ and FV0 equations, it could be expected that for a number of problems the effects of the spin can be seen directly by solving these equations in parallel.

In the general case of a Coulomb potential, any wavefunction derived using the Feshbach–Villars linearization procedure (equations (2), (24), (46), (50)–(51)) which involves a minimally coupled time derivative will have a r^{-1} term in the wavefunction, which diverges as $r \to 0$. However, when physical expectation values are considered using equations (60) and (70) then it turns out that the wavefunctions are not only normalizable but moreover the expectation value of the Coulomb energy (involving r^{-1}) is finite.

11. Concluding remarks

The theory of quantum mechanics is not only the Schrödinger equation, but the equation together with an associated mathematical formalism and physical interpretation. However, this is only an approximation and, for many problems, an equation compatible with the special theory of relativity is required. In the spin-$\frac{1}{2}$ case, the relativistic equation widely used is the Dirac equation. This does not mean, however, that *a priori* there does not exist another spin-$\frac{1}{2}$ relativistic equation. This paper has shown that an alternative derived in [1] (the FV$\frac{1}{2}$ equation) has two important features which distinguish it from the Dirac equation. These are an enlarged solution space and the requirement that an indefinite inner product is used in the associated quantum mechanical formalism.

That an alternative to the Dirac equation could mathematically exist was explained by the use of spinor representations of the proper Lorentz group. The construction of manifestly covariant equations using these quantities showed that, while the Dirac equation is the natural first-order equation, there also exists a second-order equation, denoted here as the KG$\frac{1}{2}$ equation. The KG$\frac{1}{2}$ equation contains solutions not belonging to the Dirac equation. The KG$\frac{1}{2}$ equation is, however, not directly compatible with the formalism of quantum mechanics, which requires an equation in Hamiltonian form together with a suitable inner product. Instead, the method to obtain the FV$\frac{1}{2}$ equation from the KG$\frac{1}{2}$ equation not only achieves the Hamiltonian form, but is the standard mathematical procedure to write the (indefinite) inner product in the most convenient way possible. The relativistic covariance is preserved in the FV$\frac{1}{2}$ equation, albeit not in a manifest form.

While the FV$\frac{1}{2}$ equation has eight components, it has the useful property that in the Weyl representation of the gamma matrices it decouples into two four-component equations. Only one of these needs to be solved for problems of a single particle moving in a (minimally coupled) classical external electromagnetic field, and from this solution the full eight-component wavefunctions can be easily obtained. In the case of hydrogenic atoms, it was found that the solution method was of roughly the same difficulty as the Dirac method. Also, the bound-state eight-component wavefunctions were shown to be consistent with the literature and the mathematical relationship between these and the Dirac wavefunctions established. The energy spectra and angular momentum interpretation of the states are retained. As a side benefit, the spin-0 solution can be obtained with only minor modifications. These results show that the FV$\frac{1}{2}$ equation retains the well-established description that the Dirac equation provides for hydrogenic atoms. However, this only occurs for the part of the solution space where $\langle \Psi'_{\mathrm{FV}1/2} | \Psi'_{\mathrm{FV}1/2} \rangle > 0$. The other half of the solution space is where the two distinguishing features of the FV$\frac{1}{2}$ equation manifest themselves.

While the mathematics to investigate the other half of the solution space using the indefinite inner product is provided in sections 7 and 9, the actual application to hydrogenic atoms deserves a detailed discussion and hence will be presented in a future paper. The general viewpoint taken in the present paper is that it is necessary to first establish the mathematical foundations of the FV$\frac{1}{2}$ equation.

This is the second of a series of papers. In the third the physical interpretation of the full solution space using the indefinite inner product is discussed. Thereafter it will be shown that the FV$\frac{1}{2}$ equation together with its indefinite inner product differs from other developments in the literature based upon the use of the KG$\frac{1}{2}$ equation.

Acknowledgments

The author wishes to thank Dr J M C F Govaerts and Dr B A Robson for many invaluable discussions and insights into this work. He also wishes to thank J P Costella, R Dreizler and B H J McKellar for useful comments, the Australian Government for support via an Australian Postgraduate Research Award and A-M Mårtensson-Pendrill and the Swedish Institute for financial support.

References

[1] Robson B A and Staudte D S 1996 *J. Phys. A: Math. Gen.* **29** 157

[2] Dirac P A M 1928 *Proc. R. Soc.* A **117** 610

[3] Feshbach H and Villars F 1958 *Rev. Mod. Phys.* **30** 24. Equation (2.15) in their paper is the two-component spin-0 equation that we denote as 'the Feshbach–Villars equation', or 'the FV0 equation' for convenience. In discussing this equation, authors generally refer to the paper of Feshbach and Villars, despite the fact that its content is largely based upon earlier material, in particular Taketani M and Sakata S 1940 *Proc. Math. Phys. Soc. Japan* **22** 757 and Heitler W 1943 *Proc. Roy. Irish. Acad.* **49** 1. Taketani and Sakata were the first to write Kemmer's (Kemmer N 1939 *Proc. R. Soc.* A **173** 91) ten-component first-order equation for spin-1 particles as a six-component equation in Hamiltonian form. The two-component spin-0 equation (for free particles) is easily deduced from the six-component equation, as shown by Heitler. Heitler also gives an extensive discussion of the spin-0 equation, much of which is used by Feshbach and 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. This method of derivation is referred to as 'the Feshbach–Villars linearization procedure'.

[4] Petroni N C, Gueret P, Vigier J P and Kyprianidis A 1986 *Phys. Rev.* D **33** 1674 (equation (1.2)) and references therein. Feynman R P and Gell-Mann M 1958 *Phys. Rev.* **109** 193 consider especially one of the decoupled parts of the equation in the Weyl representation of the gamma matrices, equation (6) in their paper. The full equation is given by equation (3) in their paper.

[5] Lifshitz E M, Berestetskiĭ V B and Pitaevskiĭ L P 1974 *Relativistic Quantum Theory* (Oxford: Pergamon) ch 3

[6] The notation used in this series of papers is the following. $g^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, $\overline{h} = 1$, $c = 1$. $p_\mu = (E, \boldsymbol{p})$ is the canonical 4-momentum, while $\pi_\mu = (\pi_0, \boldsymbol{\pi})$ is the mechanical 4-momentum. $A_\mu = (A_0, \boldsymbol{A})$ is the electromagnetic 4-potential, $\pi_\mu = p_\mu - eA_\mu$. $\partial_\mu = (\partial_0, \boldsymbol{\partial}) = (\frac{\partial}{\partial t}, -\nabla)$. $p_\mu \to \mathrm{i}\partial_\mu, \pi_\mu \to \mathrm{i}D_\mu$ is the quantization process. $D_\mu = \partial_\mu + \mathrm{i}eA_\mu$ is the minimal coupling. $\boldsymbol{D} = \boldsymbol{\partial} + \mathrm{i}e\boldsymbol{A}$. $a \cdot b = a_\mu b^\mu$ is the Lorentz scalar product. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, $\boldsymbol{E} = -\nabla A_0 - (\partial \boldsymbol{A}/\partial t)$, $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. The gamma matrices are defined as follows. $\gamma_\mu = (\gamma_0, \boldsymbol{\gamma})$, $(\gamma_0 = \gamma^0)$. In the standard representation,

$$\gamma^0_{\mathrm{STD}} = \begin{pmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{pmatrix} \qquad \boldsymbol{\gamma}_{\mathrm{STD}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & \mathbf{0} \end{pmatrix} \qquad \gamma_{5\mathrm{STD}} = \begin{pmatrix} \mathbf{0} & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} \end{pmatrix}$$

$$\gamma^\mu_{\mathrm{Weyl}} = U\gamma^\mu_{\mathrm{STD}}U^\dagger \qquad U = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{1}_2 & \mathbf{1}_2 \\ \mathbf{1}_2 & -\mathbf{1}_2 \end{pmatrix}$$

$\boldsymbol{\sigma}$ and τ_i are both the conventional Pauli matrices and $\not{D} = \gamma^\mu D_\mu$. The spinor quantities are defined exactly as those in chapter 3 of [5]. $[A^{\alpha\dot{\beta}}] = A_0\mathbf{1}_2 + \boldsymbol{\sigma}\cdot\boldsymbol{A}$. $A^{\alpha\dot{\beta}}$ is the spinor equivalent of the electromagnetic 4-potential A_μ. The inner product notation is similar to that used in [10]. For a given matrix $\boldsymbol{M}$, $\boldsymbol{M}^*$ means complex conjugate, $\boldsymbol{M}^T$ transpose and $\boldsymbol{M}^\dagger = \boldsymbol{M}^{*T}$. In some cases, a column vector consisting of two separate column vectors

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix}$$

is written for simplicity as $\boldsymbol{V} = (\boldsymbol{a}, \boldsymbol{b})^T$

[7] Petroni N C, Gueret P, Vigier J P and Kyprianidis A 1985 *Phys. Rev.* D **31** 3157
[8] Feynman R P 1961 *Quantum Electrodynamics* equation (9.4)
[9] Actually an extension of $\mathcal{H}$ should eventually be considered. See for example Capri A Z 1985 *Non-relativistic Quantum Mechanics* (Menlo Park: Benjamin-Cummings)
[10] Pease M C 1965 *Matrix Algebra* (New York: Academic)
[11] Greiner W 1989 *Quantum Mechanics: An Introduction* (Berlin: Springer)
[12] Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
[13] In fact both of the KG0 and KG$\frac{1}{2}$ equations are partial differential equations of the second order in each of the four variables t and $\boldsymbol{x}$. The Feshbach–Villars linearization procedure is concerned with the mathematical reduction of just the time derivative to first order, as the Hamiltonian form in quantum mechanics only requires this. Alternatives (especially Petroni N C, Gueret P and Vigier J P 1988 *Found. Phys.* **18** 1057) will be discussed in a future paper.
[14] Greiner W 1990 *Relativistic Quantum Mechanics* (Berlin: Springer)
[15] Bethe H A and Salpeter E E 1977 *Quantum Mechanics of One and Two Electron Atoms* (New York: Plenum)
[16] Auvil P R and Brown L M 1975 *Am. J. Phys.* **46** 679
[17] Davis L 1939 *Phys. Rev.* **56** 186
[18] Slater L J 1960 *Confluent Hypergeometric Functions* (Cambridge: Cambridge University Press). Consider the standard representation of the gamma matrices [6]. The operation of $(2m)^{-1}(\mathrm{i}\not{D} + m\mathbf{1}_4)$ on

$$\Psi_{\mathrm{KG}1/2} = \frac{h(r)}{r}\sqrt{2}C_1\begin{pmatrix} \mathcal{Y}^m_{lj} \\ \mathrm{i}\kappa\mathcal{Y}^m_{l'j} \end{pmatrix}$$

has the net effect of leaving the angular eigenfunctions as

$$\begin{pmatrix} \mathcal{Y}^m_{lj} \\ \mathrm{i}\mathcal{Y}^m_{l'j} \end{pmatrix}$$

and changing

$$\frac{h(r)}{r}\sqrt{2}C_1\begin{pmatrix} 1 \\ \kappa \end{pmatrix}$$

to $(g(r), f(r))^T$ given by equations (72) to (74). The action of $\boldsymbol{\sigma}\cdot\nabla$ on $R(r)\mathcal{Y}^m_{l'j}$ is given according to [5], p 108, where $R(r)$ is an arbitrary function of r. Then the linear combinations and derivatives of confluent hyergeometric functions can be simplified using results given in Slater, pp 15 and 19.