and Breach, London, to be published).

[7]We use the convention

$$\Delta_F(x^2, m^2) = i\,\langle 0|T[\phi(x)\phi(0)]|0\rangle = \frac{1}{(2\pi)^4}\int d^4p\, \frac{e^{ip\cdot x}}{m^2-p^2-i\epsilon}\,.$$

[8]See Efimov, Ref. 3, first citation.

[9]Standard textbook reference on summation methods are: E. Borel, *Leçons sur les séries divergentes* (Paris, 1928); H. S. Wall, *Continued Fractions* (New York, 1947); G. H. Hardy, *Divergent Series* (Oxford Univ. Press, Oxford, England, 1948). For applications in field theory see again Simon, Ref. 6.

[10]S. Graffi, V. Grecchi, and G. Turchetti, Nuovo Cimento 4B, 313 (1971).

[11]A. Chodos and K. Lane, Phys. Rev. D 4, 1667 (1971).

[12]H. M. Fried, Nuovo Cimento 52, 1333 (1967); Phys. Rev. 174, 1725 (1968).

[13]S. Fels, Phys. Rev. D 1, 2370 (1970).

PHYSICAL REVIEW D VOLUME 6, NUMBER 2 15 JULY 1972

Two-Component Alternative to Dirac's Equation*

L. C. Biedenharn and M. Y. Han
Duke University, Durham, North Carolina 27706

and

H. van Dam
University of North Carolina, Chapel Hill, North Carolina 27514

(Received 9 December 1971)

An alternative to Dirac's factorization of the Klein-Gordon equation is demonstrated which yields two-component, $m \neq 0$, equations. Explicit two-component solutions for the Coulomb field, $\alpha Z/r$, are given in detail; the bound energy levels are found to be precisely the same as for the Dirac equation, differences in the spectrum are discussed. The case for general electromagnetic fields is also discussed. The Poincaré invariance of this alternative equation is proved in two distinct ways: (1) by the standard method using Poincaré generators, and (2) by a group-theoretic analysis based upon Wigner's classic work. The existence of an invertible 1-1 mapping of Dirac's equation for general electromagnetic fields into and onto the alternative equation is demonstrated. That the two alternatives are not necessarily identical (for example, if chirality has a fixed significance) is discussed.

I. INTRODUCTION

The purpose of the present paper is to demonstrate that Dirac's classic derivation[1] of his famous equation for massive, charged, spin-$\frac{1}{2}$ particles admits of an alternative, which leads to a distinct, and quite remarkable, new equation. The two alternatives – Dirac's equation (labeled hereafter A_1) and the new equation (A_2) – appear as mutually exclusive alternatives in the factorization of the Klein-Gordon equation[2]; we call this the "Dirac dichotomy" since it appears already in Dirac's marvelously clear paper.[1] (We discuss this in detail in Sec. II, below.) The development of equation A_2, including the proof of Poincaré invariance, has been given in three brief – and not overly accessible – letters.[3-5]

The development of the properties of Dirac's equation has led to an enormous literature reaching over the past 43 years. Accordingly any claim to a fundamentally novel development must be treated with great skepticism. The present paper is therefore much more explicit and more detailed than is the current standard in the field. Such explicitness – including overly detailed, even repetitive, demonstrations – is especially necessary because the subject is intrinsically subtle and often (seemingly) paradoxical. Our aim is to convince the average reader, and not experts alone, as to the correctness of our results.

Let us sketch the plan of this paper. In Sec. II, we repeat Dirac's derivation of A_1, set up the dichotomy, and derive A_2 – but in a preliminary fashion, the final version of A_2 being given in Sec. IV. In Sec. III, we apply the preliminary version of A_2 to an exactly solvable problem: the Coulomb field $\alpha Z/r$.

This explicit and detailed solution is intended to show, by example, that alternative 2 *does indeed possess two-component solutions to standard problems in a manner quite impossible for alternative 1.*

In Sec. IV we return to the more abstract question of how to extend our preliminary equation so as to allow coupling to an arbitrary electromagnetic field and to allow invariance for Poincaré transformations.

Section V approaches the problems from an entirely different point of view and develops both A_1 and A_2 directly from Wigner's celebrated discussion of the unitary irreducible representations ("unirreps") of the Poincaré group $\mathcal{P}$. This rederivation of A_2 by totally distinct techniques, and concepts, is very reassuring evidence of the correctness of our alternative.

Section VI discusses in detail the particular properties possessed by A_2, and the interrelationship between the two alternatives. It is shown that, in the presence of an arbitrary electromagnetic field, there is an isomorphism between A_1 and A_2. The isomorphism does ensure that A_1 and A_2 give identical results for a spin-$\frac{1}{2}$ particle interacting with an electromagnetic field. However, the isomorphism is of a mathematical nature, and it need not be true that the physics of our equation (in the presence of weak interactions for instance) is the same as that of the Dirac equation. This is discussed further in the concluding section (Sec. VII).

Finally, to be precise let us note explicitly that our entire discussion is at the level of the one-particle Dirac equation. It is well known that, for consistency, one requires field quantization. Since quantization "cures" the two shortcomings of the one-particle Dirac theory (positive definite charge and indefinite energy) – shortcomings which A_2 also shares – we may reasonably expect a similar cure by quantizing A_2. This, however, is a task for future work and is not considered here. We will be quite happy if we convince our colleagues merely that A_2 exists and has the properties we claim.

II. THE DIRAC DICHOTOMY

Let us follow Dirac's method in developing his relativistic wave equation. In this procedure, the starting point is the relativistic invariance of the length of the four-momentum vector: $p \equiv (p_0, \vec{p})$; that is, one asserts the Klein-Gordon operator equation

$$\Box\psi(x) \equiv (p_0{}^2 - \vec{p}^2)\psi(x) = m^2\psi(x)\,. \tag{1}$$

For reasons understood rather more fully long after Dirac's initial work, one seeks to replace Eq. (1) by equations linear in all components of p. Algebraically, the problem is to factorize a five-term quadratic of the form

$$A^2 + B^2 + C^2 + D^2 + E^2 = 0\,. \tag{2}$$

This cannot be done over the field of complex numbers, but can easily be done by adjoining the elements of an appropriate Clifford algebra. The most familiar example of a Clifford algebra, the Pauli spinors (complex quaternions), allows one to factorize at most a four-term quadratic. It is thoroughly well known that one must go to the Dirac algebra to factorize Eq. (1) and/or Eq. (2). In Dirac's original notation (which is very convenient for our later discussions) one introduces two independent (= commuting) sets of Pauli spinors, $\vec{\sigma}$ and $\vec{\rho}$, and takes the outer product.

The five anticommuting units are then

$$\rho_1\sigma_i \equiv \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \text{for } i = 1, 2, 3\,,$$

$$\rho_2 \equiv \begin{pmatrix} 0 & -i1 \\ i1 & 0 \end{pmatrix},$$

and (3)

$$\rho_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Hence Eq. (2) takes a factorized form, say,

$$(A\rho_1\sigma_1 + B\rho_1\sigma_2 + C\rho_1\sigma_3 + D\rho_2 + E\rho_3)^2 = A^2 + B^2 + C^2 + D^2 + E^2\,. \tag{2'}$$

Conventionally one suppresses the fifth unit and writes Eq. (1) as

$$(p_0 + \vec{\alpha}\cdot\vec{p} + \beta m)(p_0 - \vec{\alpha}\cdot\vec{p} - \beta m)\psi = (p_0 - \vec{\alpha}\cdot\vec{p} - \beta m)(p_0 + \vec{\alpha}\cdot\vec{p} + \beta m)\psi = 0\,, \tag{4}$$

where

$$\vec{\alpha} \equiv \rho_1\vec{\sigma}, \quad \beta \equiv \rho_3\,.$$

Hence the linear equation

$$(p_0 - \vec{\alpha}\cdot\vec{p} - \beta m)\psi = 0 \tag{5a}$$

and its conjugate

$$(p_0 + \vec{\alpha}\cdot\vec{p} + \beta m)\psi = 0 \tag{5b}$$

together imply Eq. (1).

This is all very familiar without a doubt; but how unique is it? The Dirac-Clifford algebra is well known to be unique to within unitary equivalence; there is no freedom here.

That a *four-component algebra* (i.e., both $\vec{\rho}$ and $\vec{\sigma}$) *is also necessary is commonly believed. This belief we will now show to be false.*

Since it will be essential for later work (Sec. III), let us digress at this point to develop an algebra[6] which we shall designate as the algebra of "rotationally invariant Pauli spinors."

First, note that the Pauli spinors suffice to factorize the orbital angular momentum equation. That is, the four-term quadratic

$$\vec{L}^2\psi \equiv (L_x{}^2 + L_y{}^2 + L_z{}^2)\psi = l(l+1)\psi \quad (6)$$

admits the factorization

$$(\vec{\sigma}\cdot\vec{L} - l)(\vec{\sigma}\cdot\vec{L} + 1 + l)\psi = 0\,, \quad (7)$$

which is easily verified upon using the "Dirac rule":

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) = \vec{A}\cdot\vec{B} + i\vec{\sigma}\cdot\vec{A}\times\vec{B}\,, \quad (8)$$

with

$$[\vec{A},\vec{\sigma}] = [\vec{B},\vec{\sigma}] = 0\,.$$

Let us introduce the operator $\mathcal{K}$ defined by

$$\mathcal{K} = -(\vec{\sigma}\cdot\vec{L} + 1) \quad (9a)$$

and the eigenvalue equation

$$\mathcal{K}\chi_\kappa^\mu(\theta,\varphi) = \kappa\chi_\kappa^\mu(\theta,\varphi)\,. \quad (9b)$$

The (two-component spinor) eigenbasis χ_κ^μ has as sharp quantum numbers the total angular momentum j [defined by $\vec{J} = \vec{L} + \frac{1}{2}\vec{\sigma}$, $\vec{J}^2 \to j(j+1)$], sharp z component $J_z \to \mu$, and sharp orbital angular momentum l. By using the relation that $J^2 = \mathcal{K}^2 - \frac{1}{4}$, we see that $[j(\kappa) + \frac{1}{2}]^2 = \kappa^2$. Since a (nonrelativistic) spin-$\frac{1}{2}$ particle has angular momentum $j = \frac{1}{2}, \frac{3}{2}, \ldots$ [with each j occurring twice (parity)] we see that the spectrum of $\mathcal{K}$ runs over all positive and negative integers, *excluding* zero. (This last remark is very important since it shows that $\mathcal{K}^{-1}$ exists.) Defining $j(\kappa)$ to be positive yields the relation

$$j(\kappa) \equiv |\kappa| - \tfrac{1}{2}\,. \quad (10a)$$

Correlating the orbital angular momentum l [from Eq. (7)] and the parity [$\equiv (-)^l$] with the κ eigenvalue shows that

$$l(\kappa) = |\kappa| + \tfrac{1}{2}[S(\kappa) - 1] \quad (10b)$$

[where $S(\kappa)$ is defined to be the sign (±1) of κ] and

$$\text{parity} = (-)^l = (-)^{|\kappa|}S(\kappa) = (-)^{j(\kappa)+1/2}S(\kappa)\,. \quad (10c)$$

The eigenbasis $\chi_\kappa^\mu(\theta,\varphi)$ has the explicit representation

$$\chi_\kappa^\mu(\theta,\varphi) \equiv \sum_{m,\tau}\langle l(\kappa)\tfrac{1}{2}m\tau\,|\,l(\kappa)\tfrac{1}{2}j(\kappa)\mu\rangle \times Y_{l(\kappa)}^m(\theta,\varphi)\chi_{1/2}^\tau\,, \quad (11)$$

where $\langle\cdots|\cdots\rangle$ is a Wigner coefficient and $\chi_{1/2}^\tau$ denotes the basis spinors

$$\chi_{1/2}^{1/2} = \begin{pmatrix}1\\0\end{pmatrix} \quad\text{and}\quad \chi_{1/2}^{-1/2} = \begin{pmatrix}0\\1\end{pmatrix}.$$

This eigenbasis admits the pseudoscalar operator $\vec{\sigma}\cdot\hat{r}$, which is easily seen to have the property

$$\vec{\sigma}\cdot\hat{r}\chi_\kappa^\mu = -\chi_{-\kappa}^\mu\,, \quad (12)$$

$$(\vec{\sigma}\cdot\hat{r})^2 = 1\,. \quad (13)$$

It therefore follows that the operators $\mathcal{K}$ *and* $\vec{\sigma}\cdot\hat{r}$ *anticommute.*

But if there exist *two* anticommuting operators then there must exist *three* (the third being the product). Hence we have *three anticommuting rotationally invariant operators*: $\{\mathcal{K}, \sigma\cdot\hat{r}, \sigma\cdot\hat{r}\mathcal{K}\}$, with "rotational invariance" being defined precisely as invariance under the specific generator $\vec{J} \equiv \vec{L} + \frac{1}{2}\vec{\sigma}$.

If we normalize our operators to have unit square – which is always possible since $\mathcal{K}^2$ is nonvanishing, and $\vec{\sigma}\cdot\hat{r}$ is already properly normed – we see that we have our desired *rotationally invariant Pauli spinors*:

$$\eta_1 \equiv \vec{\sigma}\cdot\hat{r}\,, \quad (14a)$$

$$\eta_2 \equiv i\eta_1\eta_3\,, \quad (14b)$$

$$\eta_3 \equiv (-)^{j(\kappa)+1/2}S(\kappa)\,, \quad (14c)$$

which obey the relations

$$\eta_i\eta_j = i\epsilon_{ijk}\eta_k \;\text{ for } (ijk) = 1,2,3\,, \quad (14d)$$

$$(\eta_i)^2 = 1\,. \quad (14e)$$

[Note the curious phase given in Eq. (14c). As far as the Pauli properties Eqs. (14d) and (14e) are concerned, this choice is one of an unlimited number of equally good possibilities (the possibility of rephasing every j subspace by + or −). But in order to develop a wave equation for arbitrary electromagnetic fields, we must require η_3 to have the significance of the (internal) parity operator P_{int} acting on the particle variables $\vec{x}$ and $\vec{p}$. (This will be clearer in the sequel.) Note also that this choice was not made in Ref. 3 – where we used "η_3" $= S(\mathcal{K})$. In that paper this led to the assertion that Lorentz invariance fails.]

The operator $\vec{\sigma}\cdot\vec{p}$ has an expansion in terms of these η_i operators, namely,

$$\vec{\sigma}\cdot\vec{p} = \eta_1 p_r + (j+\tfrac{1}{2})(-)^{j-1/2}\eta_2\,, \quad (15)$$

where

$$p_r = \text{radial momentum operator} \equiv \hat{r}\cdot\vec{p} - i/r \quad (16a)$$

and

$$[p_r, \eta_i] = 0\,. \quad (16b)$$

The significance of the decomposition of $\vec{\sigma}\cdot\vec{p}$ given in Eq. (15) is that one finds

$$[\vec{\sigma}\cdot\vec{p}, \eta_3]_+ = 0\,, \quad (17)$$

a result that can also be very easily demonstrated directly.

After this digression, let us return to our problem. We can now demonstrate by explicit construction *that there exists a two-component (spinor) factorization of the Klein-Gordon equation,* namely,

$$(p_0 - \vec{\sigma}\cdot\vec{p} - \eta_3 m)(p_0 + \vec{\sigma}\cdot\vec{p} + \eta_3 m)\psi = (p_0 + \vec{\sigma}\cdot\vec{p} + \eta_3 m)(p_0 - \vec{\sigma}\cdot\vec{p} - \eta_3 m)\psi = 0\,. \tag{18}$$

[*Proof:* The operators p_0 and $\vec{\sigma}\cdot\vec{p}$ certainly commute with each other, and p_0 also commutes with η_3 (since this operator is composed from $\vec{\sigma}\cdot\vec{L}$, which itself commutes with p_0). We have already shown that $\vec{\sigma}\cdot\vec{p}$ anticommutes with η_3 in Eq. (17). Hence the assertion follows.]

Thus we have two possible candidates for a two-component first-order wave equation for massive spin-$\frac{1}{2}$ particles: either

$$p_0\psi = (\vec{\sigma}\cdot\vec{p} + \eta_3 m)\psi \tag{19a}$$

or its conjugate

$$p_0\psi = -(\vec{\sigma}\cdot\vec{p} + \eta_3 m)\psi\,. \tag{19b}$$

We shall discuss in detail in subsequent sections the properties of Eqs. (19a) and (19b).

Let us turn to a different question now: Does our construction contradict any claims as to uniqueness given in Dirac's paper? Quite surprisingly, perhaps, it does not, for Dirac – with commendable caution (and insight) – carefully phrased his argument[7]:

"The symmetry between p_0 and p_1, p_2, p_3 required by relativity shows that, since the Hamiltonian we want is linear in p_0, it must also be linear in p_1, p_2, and p_3. Our wave equation is therefore of the form

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m)\psi = 0\,, \tag{4}$$

where for the present *all that is known about the dynamical variables or operators* $\alpha_1, \alpha_2, \alpha_3, \beta$ *is that they are independent of* p_0, p_1, p_2, p_3, *i.e., that they commute with* t, x_1, x_3, x_3. *Since we are considering the case of a particle moving in empty space, so that all points in space are equivalent, we should expect the Hamiltonian not to involve* t, x_1, x_2, x_3. *This means that* $\alpha_1, \alpha_2, \alpha_3, \beta$ *are independent of* t, x_1, x_2, x_3, *i.e., that they commute with* p_0, p_1, p_2, p_3. *We are therefore obliged to have other dynamical variables besides the co-ordinates and momenta of the electron, in order that* $\alpha_1, \alpha_2, \alpha_3, \beta$ *may be functions of them.* The wave function ψ must then involve more variables than merely x_1, x_2, x_3, t."

Thus we see at once what it is that makes our factorization work: *The factorizing matrices which we have introduced* – in particular, η_3 – *explicitly depend on space-time, and what is more, they depend on a specific point in space-time that has been singled out* (the origin of our spherical basis). *Our two-component factorization,* then, *in no way contradicts Dirac's work.*

[Equation (19) is preliminary and needs to be changed (or more accurately reinterpreted) for two reasons. The first reason is that Eq. (19) is not invariant for Poincaré transformations; it is not even invariant for spatial translations. The second reason, which is related to the first, is that Eq. (19) cannot readily incorporate interactions with a general electromagnetic field. The only change which is needed is to extend the significance of η_3 in Eq. (19) so that it includes "external parity," a notion which we shall define in the next sections. To convince the reader that this change is sufficient we shall proceed in a cautious manner. First, in Sec. III, we treat in detail the case of the Coulomb field with a center which coincides with the center for η_3. This special electromagnetic field is easily incorporated in Eq. (19). Consideration of the formally equivalent case where the center of the Coulomb field does not coincide with that of η_3 already suggests the needed modification of Eq. (19). In the first part of Sec. IV we study the change needed for our equation so that it can incorporate the interaction with an arbitrary electromagnetic field. In this we are guided by the demand that our equation must factorize the Kramers equation for an arbitrary electromagnetic field. In the last part of Sec. IV we establish that the new equation is invariant for Poincaré transformations (≡proper orthochronous inhomogeneous Lorentz transformations).]

Let us summarize the results obtained in this section as a formal dichotomy, which we term the "Dirac dichotomy": *The factorization of the Klein-Gordon equation into two commuting linear operators in p allows two mutually exclusive alternatives:*

either A_1: the factorizing matrices are independent of space-time,

or A_2: the factorizing matrices are *not* independent of space-time.

A_1 leads to Dirac's equation, whereas A_2 leads to Eq. (19). In contrast to Dirac's equation, Eq. (19) is not manifestly Poincaré-invariant; however in succeeding sections we shall find an interpretation of (19) which is indeed Poincaré-invariant.

III. EXPLICIT TWO-COMPONENT SOLUTIONS FOR THE COULOMB FIELD $\alpha Z/r$

The classic example for testing the applicability of spin-$\frac{1}{2}$ wave equations is to solve the "hydrogenatom problem," that is, to solve the problem of a pure Coulomb field interaction, $\alpha Z/r$, where r measures (presumably) the distance of the spin-$\frac{1}{2}$ particle from a rigidly fixed source having for simplicity no other attributes. For the Dirac equation this problem, as is well known, was solved perturbatively in Dirac's definitive paper[1]; exact solutions were found shortly after,[8] and have been a continuing source of physical insight (and esoterica) ever since.[9]

The equation which we wish to solve is this:

$$H\psi(\vec{r}) \equiv \left(\frac{\alpha Z}{r} + \vec{\sigma}\cdot\vec{p} + \eta_3 m\right)\psi(\vec{r}) = E\psi(\vec{r}), \tag{20}$$

where

$$\eta_3 \equiv (-)^{j(\kappa)+1/2} S(\kappa) = P_{\text{int}}, \tag{21}$$

and κ, $S(\kappa)$, and P_{int} are defined explicitly in Eqs. (9) and (10) of Sec. II. The first step is to use Dirac's spinor form for the radial momentum, that is,

$$\vec{\sigma}\cdot\vec{p} = -i\eta_1\left(\frac{\partial}{\partial r} + \frac{1}{r}\right) - r^{-1}|\kappa|(-)^{\kappa}\eta_2. \tag{22}$$

[The form in Eq. (22) uses the definitions given in Sec. II for the rotationally invariant Pauli spin operators η_i.]

The most economical way to solve Eq. (20) would be to iterate this equation, thereby obtaining the (chirality-projected, i.e., two-component) Kramers equation with known solutions.

We shall follow the more pedestrian approach of a direct solution of Eq. (20), since in this way any possible objection as to the validity of iteration is eliminated from the beginning. Accordingly we introduce a two-component spherical basis for the wave functions $\psi(\vec{r})$ having radial functions as scalar multipliers; that is

$$\psi(\vec{r}) = \sum_{\kappa,\mu} f_\kappa(r)\chi_\kappa^\mu(\theta,\varphi), \tag{23}$$

where the spinor spherical functions χ_κ^μ have been defined in Eq. (11).

Upon introducing (22) and (23) into (20) we obtain a radial differential equation:

$$\left(\frac{\alpha Z}{r} - E + m(-)^{l(\kappa)}\right)f_\kappa + i\left(\frac{d}{dr} + \frac{1-\kappa}{r}\right)f_{-\kappa} = 0. \tag{24}$$

The essential content of (24) can be made much more evident if we rewrite it in a two component form, where the two components are distinguished by the sign of κ:

$$\begin{pmatrix} \frac{\alpha Z}{r} - E + m(-)^{j+1/2} & i\left(\frac{d}{dr} + \frac{\frac{1}{2}-j}{r}\right) \\ i\left(\frac{d}{dr} + \frac{j+\frac{3}{2}}{r}\right) & \frac{\alpha Z}{r} - E + m(-)^{j-1/2} \end{pmatrix}\begin{pmatrix} f_{(\kappa)} \\ f_{-(\kappa)} \end{pmatrix} = 0. \tag{24'}$$

The form of Eq. (24′) shows that *the two values of the* (internal) *parity function as a formal two-component space in a direct formal analogy to the two-component Dirac* $\vec{\rho}$ *space.*

Aside from the insight (parity space $\leftrightarrow$ ρ space) that Eq. (24′) yields, it has not much to recommend it, for a brute-force attempt to solve Eq. (24′) is messy and uninformative. We will therefore proceed by an alternative route, having a rather deeper physical meaning. Let us now introduce a very different spin-angle basis, which is a *unitary* transform of the original spin-angle basis χ_κ^μ. That is, we define the new basis

$$\chi_\gamma^\mu(\theta,\varphi) \equiv \left[\exp[\tfrac{1}{2}(-1)^{\kappa+1}]\tanh^{-1}\left(\frac{\alpha Z}{|\kappa|}\right)\eta_2\right]\chi_\kappa^\mu(\theta,\varphi) \tag{25a}$$

$$\equiv T\chi_\kappa^\mu(\theta,\varphi). \tag{25b}$$

This is a very complicated basis to introduce "out of the blue." Shortly we shall discuss the motivation which leads to this basis; for reasons discussed later on, we designate this transformation as the "Sommerfeld rotation" (even though it is not a three-space rotation at all).

The properties of this new basis are easily developed. Let us introduce a new operator

$$\Gamma \equiv \mathcal{K} - i\alpha Z\vec{\sigma}\cdot\hat{r}, \tag{26}$$

which we will call "the two-component form of Temple's operator," in exact analogy to $\mathcal{K}$ which is the two-component form of Dirac's operator $K = \rho_3\mathcal{K}$. (This designation for Γ is discussed below.)

One now establishes easily the following properties:

(a) $$\Gamma = \frac{|\gamma|}{|\kappa|}T\mathcal{K}T^{-1}, \tag{27}$$

(b) $$\Gamma\chi_\gamma^\mu = |\gamma|S(\kappa)\chi_\gamma^\mu \equiv \gamma\chi_\gamma^\mu, \tag{28}$$

where we have defined

(c) $$|\gamma| \equiv |[\kappa^2 - (\alpha Z)^2]^{1/2}|, \tag{29}$$

(d) $$\gamma = \pm|\gamma|, \quad \text{sgn}(\gamma) \equiv \text{sgn}(\kappa). \tag{30}$$

From the defining relations for the spin-angle functions $\chi_\kappa^\mu(\theta,\varphi)$ we can – using Eqs. (25)–(30) – derive the spectrum for Γ. For every allowed value of κ and μ there is precisely one function χ_γ^μ; the spectrum of Γ is therefore $\pm|n^2 - (\alpha Z)^2|^{1/2}$ where n runs over all positive nonzero integers once. (Note that the square root is always real if $\alpha Z < 1$, as we shall assume to be the case henceforth.)

Let us now carry out this same transform on Eq. (20) itself. To do so we use the results

$$T\eta_1 T^{-1} = \eta_1\left(\left|\frac{\kappa}{\gamma}\right| + (-)^\kappa \frac{\alpha Z}{|\gamma|}\eta_2\right)$$

$$= \left|\frac{\kappa}{\gamma}\right|\eta_1 + (-)^\kappa \frac{i\alpha Z}{|\gamma|}\eta_3\,, \tag{31a}$$

$$T\eta_3 T^{-1} = \left|\frac{\kappa}{\gamma}\right|\eta_3 + (-)^{\kappa+1}\frac{i\alpha Z}{|\gamma|}\eta_1\,. \tag{31b}$$

Introducing (22) into (20), multiplying through by $i\eta_1$, and transforming [thereby using (21)] yields the desired form:

$$\left[\frac{\partial}{\partial r} + \frac{1}{r}(1 + (-)^\kappa|\gamma|\eta_3) - \frac{iE(\kappa)}{|\gamma|}\eta_1 + m\eta_2 + \frac{(-)^\kappa \alpha Z E}{|\gamma|}\eta_3\right]T\psi = 0\,. \tag{32}$$

To solve this equation we introduce the new spin-angle functions; that is, from Eq. (25) we have

$$\psi' \equiv T\psi \equiv \sum_\kappa f_\kappa(r)\chi^\mu_\gamma$$

$$= \sum_{\kappa>0}\left(f^{(r)}_{(\kappa)}\chi^\mu_{(\gamma)} + f^{(r)}_{-(\kappa)}\chi^\mu_{-(\gamma)}\right). \tag{23$'$}$$

This substitution puts Eq. (32) in a two-component form:

$$\begin{pmatrix} \frac{d}{dr} + \frac{1+|\gamma|}{r} + \frac{\alpha Z E}{|\gamma|} & \frac{iE|\kappa|}{|\gamma|} + im(-)^\kappa \\ \frac{iE|\kappa|}{|\gamma|} - im(-)^\kappa & \frac{d}{dr} + \frac{1-|\gamma|}{r} - \frac{\alpha Z E}{|\gamma|} \end{pmatrix}\begin{pmatrix} f_{(\gamma)} \\ f_{-(\gamma)} \end{pmatrix} = 0\,. \tag{32$'$}$$

Now it has been shown[10,11] that the radial differential equations for the Dirac-Coulomb functions take a remarkably elegant form in the coordinate frame defined by the Sommerfeld rotation. Using normalized continuum functions[12] we have the explicit definitions

$$F_{l(\gamma),\eta}(kr) \equiv C_l(\eta)(kr)^{l(\gamma)}e^{-ikr} \times {}_1F_1(l(\gamma)+1-i\eta,\, 2l(\gamma)+2,\, 2ikr)\,, \tag{33a}$$

$$C_{l(\gamma)}(\eta) = \frac{2^{l(\gamma)}e^{-\pi\eta/2}|\Gamma(l(\gamma)+1+i\eta)|}{\Gamma(2l(\gamma)+2)}\,, \tag{33b}$$

$$l(\gamma) \equiv |\gamma| + \tfrac{1}{2}[\mathrm{sgn}(\gamma) - 1]\,, \tag{33c}$$

$$k \equiv |(E^2 - m^2)^{1/2}|\,, \tag{33d}$$

and

$$\eta \equiv \alpha Z E/k = \text{relativistic Sommerfeld number.} \tag{33e}$$

When so defined, we have a concise operational definition of the $F_{l,\eta}(kr)$:

$$\left(\frac{kr\eta}{\gamma} + \frac{d}{dr} + \frac{(1+\gamma)}{r}\right)F_{l(\gamma),\eta}(kr) = k\,\mathrm{sgn}(\gamma)\,|(1+\tfrac{1}{2}\eta^2)|^{1/2}F_{l(-\gamma),\eta}(kr)\,. \tag{34}$$

Equations (33) and (34) are remarkable in that they constitute a uniform definition not only of the relativistic Dirac-Coulomb-Sommerfeld radial functions,[10] but also the nonrelativistic Coulomb radial functions ($c\to\infty$) and the plane-wave (spherical basis) radial functions

$$F_{l(\gamma),\eta=0}(kr) = j_{l(\gamma)}(kr) = \text{spherical Bessel functions}$$

for *both* the relativistic and nonrelativistic cases.

To apply (33) and (34) to (32$'$) we introduce

$$f_{|\gamma|} = AF_{l(|\gamma|),\eta}(kr)$$

and

$$f_{-|\gamma|} = BF_{l(-|\gamma|),\eta}(kr)\,. \tag{35}$$

Equation (32$'$) then implies that the homogeneous system in A and B has a nontrivial solution given by

$$\frac{A}{B} = i\left|\left(\frac{E|\kappa| + m|\gamma|(-)^\kappa}{E|\kappa| + m|\gamma|(-)^{\kappa+1}}\right)^{1/2}\right|\,. \tag{36}$$

Putting these results together, we find the desired (unnormalized) solutions to be

$$\psi' = i|(E|\kappa| + (-)^\kappa m|\gamma|)^{1/2}|F_{l(|\gamma|),\eta}(kr)\chi^\mu_{+|\gamma|}(\theta,\varphi) + |(E|\kappa| + (-)^{\kappa+1}m|\gamma|)^{1/2}| \times F_{l(-|\gamma|),\eta}(kr)\chi^\mu_{-|\gamma|}(\theta,\varphi)\,. \tag{37}$$

(Transforming back to the original frame simply replaces the spin-angle functions above by $\chi^\mu_{\pm|\kappa|}$.)

Equation (37) constitutes the complete two-component solution to the relativistic Coulomb problem for alternative 2 in all (three) cases: continuum, zero-energy "continuum," and (by analytic continuation) the discrete (bound) solutions.

Let us consider the discrete set of solutions further. Using Eq. (33a), it is clear from the properties of the confluent Gauss function ${}_1F_1$ that the infinite series terminates for the condition $l(\gamma)+1-i\eta$ = nonpositive integer. This condition is more critical for γ positive. For this case we have $|\gamma| - i\eta = -(N - |\kappa|)$, where (in agreement with the ordering of the Bohr spectrum) N is the principle quantum number, $N \geq 1$.

Equation (37) shows that *for each admissable N and $|\kappa|$ there is, at most, one solution.* The energy condition $|\gamma| - i\eta = -(N - |\kappa|)$ allows a solution for $|\kappa| = 1, 2, \ldots, N$. For the level N, $|\kappa| = N$, however, we find that $|\gamma| - i\eta = 0$; hence there is a terminating solution $F_{l(\gamma),\eta}$ only for γ negative. If we examine Eq. (37), however, we see that for $|\kappa| = N =$ even integer we find ψ' vanishes, since only $F_{l(-\gamma),\eta}$ exists and its coefficient in (37) is zero. For $|\kappa| = N =$ odd integer the wave function ψ' exists (but note that it has a sharp parity).

The situation is the reverse for the conjugate A_2 equation; that is, for the equation $H = (\alpha Z/r) - \vec{\sigma}\cdot\vec{p} - m\eta_3$. The change in the sign of the mass term now allows [cf. Eq. (37)] a nonvanishing solution for $|\kappa| = N =$ even integer but no solutions for $|\kappa| = N =$ odd integer.

Hence we have shown the following results:

Result (a):

$$H = \frac{\alpha Z}{r} + \vec{\sigma}\cdot\vec{p} + m\eta_3 .$$

The energy levels are precisely the Dirac-Sommerfeld levels *but each level is nondegenerate* and solutions for $|\kappa| = N =$ even integer do not exist.

Result (b):

$$H = \frac{\alpha Z}{r} - \vec{\sigma}\cdot\vec{p} - m\eta_3 .$$

The energy levels are again precisely the Dirac energy levels but once again the levels are nondegenerate and all levels having $|\kappa| = N =$ odd integer do not occur.

In other words the two two-component Hamiltonians $H_\pm \equiv (\alpha Z/r) \pm (\vec{\sigma}\cdot\vec{p} + m\eta_3)$ divide up the Dirac-Coulomb spectrum in such a way as to produce two nondegenerate spectra. For the doubly degenerate Dirac levels ($|\kappa| \neq N$) each two-component Hamiltonian has one of the two degenerate levels, for the nondegenerate Dirac levels ($\kappa = -N$), the levels belong alternately to the two Hamiltonians.[13]

In order to demonstrate how closely the two-component solutions, developed above, are related to the Dirac-Coulomb solutions let us now make use of the iterated equation.[10] The iterated Dirac-Coulomb equation is

$$(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{p} - \rho_3 m)(\pi_0 + \rho_1\vec{\sigma}\cdot\vec{p} + \rho_3 m)\psi = 0$$

or

$$\left\{\left(E - \frac{\alpha Z}{r}\right)^2 + \rho_1\left[\vec{\sigma}\cdot\vec{p}, \frac{\alpha Z}{r}\right] - \vec{p}^2 - m^2\right\}\psi = 0 ,$$
$$\left[\left(E - \frac{\alpha Z}{r}\right)^2 + \frac{i\alpha Z}{r^2}\rho_1\vec{\sigma}\cdot\hat{r} - \vec{p}^2 - m^2\right]\psi = 0 . \tag{38}$$

Introducing a spherical basis, one has first

$$\vec{p}^2 = -\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{\vec{L}^2}{r^2} ,$$

and secondly the spinor factorization of $\vec{L}^2$: $\vec{L}^2 = \mathcal{K}(\mathcal{K}+1)$ so that (38) becomes

$$\left[\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + (E^2 - m^2) - \frac{2\alpha ZE}{r} + \frac{1}{r^2}(\mathcal{K}(\mathcal{K}+1) - (\alpha Z)^2 + i\alpha Z\rho_1\vec{\sigma}\cdot\hat{r})\right]\psi = 0 . \tag{39}$$

The terms multiplying $1/r^2$ can be given an interesting form. Define the operator

$$\tilde{\Gamma} \equiv \mathcal{K} + i\alpha Z\rho_1\vec{\sigma}\cdot\hat{r} . \tag{40a}$$

Then it is easily found that the bracketed terms $(\cdots)$ in (39) take the form

$$\mathcal{K}(\mathcal{K}+1) - (\alpha Z)^2 + i\alpha Z\rho_1\vec{\sigma}\cdot\hat{r} = \tilde{\Gamma}(\tilde{\Gamma}+1); \tag{40b}$$

hence if $\tilde{\Gamma}$ is brought to diagonal form, $\tilde{\Gamma} \to \gamma = \pm[\kappa^2 - (\alpha Z)^2]^{1/2}$, Eq. (39) becomes a purely radial equation with no spin or ρ matrices:

$$\left(\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr} + k^2 - \frac{2k\eta}{r} + \frac{l(\gamma)(l(\gamma)+1)}{r^2}\right)F(r) = 0 . \tag{41}$$

[The solutions $F(r)$ are just those given earlier in (33).]

The operator $\tilde{\Gamma}$ was first introduced by Temple[14] in 1934; it is, however, not a constant of the motion. Temple's operator has been rediscovered many times since.[15] The transformation which diagonalizes $\tilde{\Gamma}$ has been developed in Ref. 10:

$$S = \exp\left\{\tanh^{-1}\frac{\alpha Z}{|\kappa|}\rho_2\vec{\sigma}\cdot\hat{r}K\right\} , \tag{42}$$

where $K = \rho_3\mathcal{K}$ is Dirac's operator. (Note the remarkable fact that the angular momentum is invariant under S.)

In Ref. 10 it is shown that, *in the classical limit,* S transformations correspond to the rotation by which Sommerfeld[16] derived the "Dirac-Coulomb" energy levels in 1916; hence the name "Sommerfeld's rotation" for S.

Consider next the Kramers equation for the Coulomb field; this is simply Eq. (39) with $\rho_1 \to -1$ (say). Projecting on ρ_1 (chirality) also changes Temple's operator to the two-component form Γ, given in Eq. (26). Once again there is a transformation, T in Eq. (25), which diagonalizes Γ.

Let us summarize by noting the following significant features upon comparing the Dirac-Coulomb problem and the two-component problem:

(a) The two-component solutions are linear combinations having terms of both parities; *for the special case* $|\kappa| = N$, *however, the parity is sharp.*

(b) The Dirac-Coulomb solutions are linear combinations having terms in both basis states of Dirac (ρ) space; *for the special case* $|\kappa| = N$, *however,* ρ_3 *becomes sharp.*[11]

(c) The transformation that diagonalizes the two-component form of Temple's operator mixes parity.

(d) The transformation that diagonalizes Temple's operator $\vec{\Gamma}$ mixes ρ_3 components.

It is clear that the wave functions of the two-component versus four-component problem (A_1 vs A_2) are very closely related. In fact, *one sees that the wave functions for the two-component problem can be "assembled" by adding together the* $\rho_3 \to +1$ *and* $\rho_3 \to -1$ *projections of the Dirac-Coulomb wave function,* where the latter is in the transformed (T-frame) coordinate system.

One further point is to be noted: The transformation T is a *symmetry operation* and leaves the Dirac inner product $\langle\bar{\psi}|\psi\rangle \equiv \langle\psi^\dagger\rho_3|\psi\rangle$ invariant. By analogy, $\rho_3 \to \eta_3$, one expects the inner product

$$\langle\bar{\varphi}|\varphi\rangle \equiv \langle\varphi^\dagger\eta_3|\varphi\rangle$$

to be invariant under T; that this is indeed correct is readily verified. (This inner product is discussed further in Sec. IV B.)

It will be noted that we have chosen a special coordinate frame (whose origin is defined by the nucleus) in which to solve Eq. (20). What would happen[17] if, purely formally, one were to choose a coordinate system for calculational purposes, displaced by a distance $\vec{a}$, say, from the nucleus? Surely if there is any physical sense to Eq. (20) such a formal change must be of no consequence. In the new origin, η_3 represents the (internal) parity operator (since it is defined in terms of $\vec{\sigma}$ and $\vec{L}$ and the latter refers to the new origin). But the interaction now appears very different, becoming $\alpha Z/|\vec{r}-\vec{a}|$. More disconcerting, however, is the fact that *the operator* η_3 *no longer commutes with the interaction.*

This latter result is *totally unacceptable*; among other things it spoils the iteration of Eq. (20) into the Kramers equation.

The way to retrieve the situation is probably clear: The operator η_3, which anticommutes with $\vec{r}$, *must somehow also anticommute with* $\vec{a}$. Recognizing that $\vec{a}$ is a constant polar vector (a fixed displacement), we see that we must adjoin to $\eta_3 = P_{\rm int}$, the operator $P_{\rm ext}$ which is defined to be an *external parity operator,* and hence, by definition, the new η_3 anticommutes with objects, such as $\vec{a}$, which have odd parity but which (being independent of $\vec{r},\vec{p}$) commute with $P_{\rm int}$. We show in detail in Sec. IV how this extension neatly retrieves the situation.

[Let us note here that there exists, for A_2, a second exactly solvable problem, the problem of the constant magnetic field: $A_0=0$, $\vec{A}=\frac{1}{2}(\vec{B}_0\times\vec{r})$. This problem played an important role in the development of A_2 in that it clearly showed the necessity for η_3 to be the (internal) parity operator. For the purposes of the present paper, however, this problem plays a subsidiary role, and it seems reasonable to develop these explicit two-component solutions elsewhere.]

IV. FURTHER DEVELOPMENTS

A. The Factorization of the Kramers Equation

The question as to why the Dirac equation has four – and not just two – components has been a continuing source of inquiry ever since Dirac's original paper became the accepted basis for the theory of spin-$\frac{1}{2}$ particles. Feynman,[18] in particular, has discussed this question from many angles. Let us begin here with the approach given by Feynman and Gell-Mann in their classic paper on chiral currents in weak interactions.[19]

The Dirac equation in an arbitrary external electromagnetic field [using the canonical substitution: $p \to \pi \equiv p - eA$ in Eq. (5)] does not – upon iteration – obey the Klein-Gordon equation having p replaced by π. Instead, one obtains, as is well known, the equation

$$[\pi\cdot\pi + e\vec{\sigma}\cdot(i\rho_1\vec{E}+\vec{B}) - m^2]\psi = 0\,. \qquad (43)$$

Since Eq. (43) contains but the single ρ matrix ρ_1, it can be split (*Poincaré-invariantly*) into two separate second-order two-component equations:

$$\rho_1 \to +:\quad \pi_+\pi_-\psi = m^2\psi\,, \qquad (44a)$$

$$\rho_1 \to -:\quad \pi_-\pi_+\psi = m^2\psi\,, \qquad (44b)$$

where

$$\pi_\pm \equiv \pi_0 \pm \vec{\sigma}\cdot\vec{\pi}\,. \qquad (44c)$$

These two-component equations had been otained much earlier – and in a different way – by van der Waerden[20] and by Kramers.[21]

Note especially that Eq. (43) – split into the two "chiral" forms of the so-called Kramers equation [Eqs. (44a) and (44b)] – does not have sharp parity or sharp C, but only the product CP.[22] Note, also, that the electromagnetic fields enter in the chiral, or Silberstein, form: $\pm i\vec{E}+\vec{B}$.

The Kramers equation [either (44a) or (44b)] has some remarkable properties, as has been discussed in detail by Brown in his Colorado lectures.[23] In particular, a generalized form of the Feynman rules can be set up[24] so as to yield an acceptable quantum electrodynamics, but curiously – as Brown[23] has shown – the Kramers equation is inconsistent unless the gyromagnetic ratio has precisely the numerical value $g=2$.

It was by noting the chiral (and CP) symmetry evidenced by the Kramers equations that Feynman and Gell-Mann motivated their postulate that the weak-interaction current is to be a chirally projected Dirac current. [This, of course, necessarily denies that the Dirac Hamiltonian generates the time displacements for the weak current. However, the really essential conclusions for weak interactions depend primarily on the *form* of the weak current (kinematics); in fact, Hamiltonian considerations are largely irrelevant in a perturbation theory that stops at first order anyway. We discuss this further in Sec. VII.]

We shall now demonstrate that either, and hence both, of Kramers's two equations may be factorized over the Pauli (two-component) spinors.

Proof. This claim is certainly true if there are no external fields (cf. Sec. II), or if we have either of the two cases discussed in Sec. III:

(a) $A_0 = \alpha Z/r, \quad \vec{A} = 0,$

(b) $A_0 = 0, \quad \vec{A} = \frac{1}{2}(\vec{B}_0 \times \vec{r}).$

In fact, one can easily verify that the combined case $A_0 = \alpha Z/r$, $\vec{A} = \frac{1}{2}(\vec{B}_0 \times \vec{r})$ also works out. That is, define

$$\mathcal{O}_{\pm} \equiv \eta_3(\pi_0 - \rho_1 \vec{\sigma}\cdot\vec{\pi}) \pm m, \tag{45}$$

where

$$\pi_0 = p_0 - \alpha Z/r,$$

$$\vec{\pi} = \vec{p} - \tfrac{1}{2}e(\vec{B}_0 \times \vec{r}),$$

$$\vec{B}_0 = \text{constant axial vector.}$$

Then

$$\begin{aligned}\mathcal{O}_+\mathcal{O}_- = \mathcal{O}_-\mathcal{O}_+ &= \eta_3(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{\pi})\eta_3(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{\pi}) - m^2 \\ &= (\pi_0 + \rho_1\vec{\sigma}\cdot\vec{\pi})(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{\pi}) - m^2 \\ &= \pi_0{}^2 + \rho_1[\vec{\sigma}\cdot\vec{\pi}, \pi_0] - (\vec{\sigma}\cdot\vec{\pi})^2 - m^2 \\ &= \pi\cdot\pi - m^2 + i\alpha Z\rho_1\vec{\sigma}\cdot\vec{\nabla}(1/r) + i\vec{\sigma}\cdot\vec{B}_0. \end{aligned}\tag{46}$$

This is precisely the iterated Dirac equation and agrees (after $\rho_1 \to \pm 1$) with the Kramers equation (44a) and (44b).

The reason why this demonstration works is easily seen: *It works because one can explicitly verify that for the fields given in Eq. (45) one has the relations*

$$[\eta_3, \pi_0] = 0 \tag{47a}$$

and

$$[\eta_3, \vec{\sigma}\cdot\vec{\pi}]_+ = 0. \tag{47b}$$

Equations (47a) and (47b) define what Good[25] calls an "even environment." Clearly the proof extends *to all electromagnetic fields having an even environment.*

If one could have Eqs. (47a) and (47b) true for *arbitrary* electromagnetic fields then the demonstration would be complete. But it will require a far-reaching generalization before we can make this conclusion. Somehow or other we must generalize to the case of an odd environment. [We will interrupt the proof at this point, and complete it beginning with Eqs. (51a) and (51b) and ending with the proposition I.]

The necessity for some sort of generalization is implicit already in the Coulomb-field example, as discussed briefly at the end of Sec. III. It is physically evident that in this problem there exists a preferred coordinate frame, namely, that frame of reference in which the source of the $\alpha Z/r$ field defines the origin – and accordingly defines the (internal) parity operator η_3. While this frame is certainly defined physically it is by no means required mathematically. We can at will choose to calculate in any fixed frame we so desire. But if we really do opt for a mathematically displaced frame (displaced by $\vec{a}$) then

$$\frac{\alpha Z}{r} \to \frac{\alpha Z}{|\vec{r} - \vec{a}|}$$

and our interaction relative to the new calculational frame *has components with both an odd and an even environment, which we do not know (as yet) how to handle.*

We are thus in an absurd situation, for no physical theory can depend in an essential way on a purely mathematical frame of reference.

Let us pursue this dilemma further. From the multipole expansion we can easily expand $|\vec{r} - \vec{a}|^{-1}$ in spherical harmonics around the new origin. (The radius $|\vec{r}| = |\vec{a}|$ will be formally singular, but since this singularity is in radial space – and hence commutes with the parity operation – we can ignore this complication as inessential.)

The essential point is that $|\vec{r} - \vec{a}|^{-1}$ will involve terms such as $Y_l^m(\hat{r})Y_l^{m*}(\hat{a})$ which have parity $(-)^l$ under η_3. We may single out the (irregular) dipole term $\vec{a}\cdot\vec{\nabla}(1/r)$ as the prototype for our dilemma.

Our problem is thus to achieve – in some rational way – the property

$$[\eta_3, \vec{a}\cdot\vec{\nabla}(1/r)] = 0. \tag{48}$$

Clearly one way to do this is to adjoin the polar vector $\vec{a}$ as a formal parity-changing element in our calculational structure. Correspondingly we must generalize η_3 (which necessarily commutes with with the constant vector $\vec{a}$) to include a formal element – call it the *external* parity operator P_{ext} – anticommuting with $\vec{a}$. That is,

$$\eta_3 \to \eta_3^{\text{extended}} \equiv \eta_3 P_{\text{ext}}. \tag{49}$$

By definition, then, we achieve our goal:

$$[\eta_3^{\text{ex}}, \vec{a}\cdot\vec{\nabla}(1/r)] = 0, \quad \eta_3^{\text{ex}} = \eta_3^{\text{extended}}. \tag{50}$$

This step is both far-reaching and – as we have found in discussions with colleagues – highly disturbing. It smacks too much of a trick, possibly even a self-deluding trick.

Let us face this unease directly, and subject our analysis to criticism:

(1) Is this step sufficient to accomplish the goal or do we have to patch things up, *ad hoc*, for every new problem?

A little reflection will show that *every scalar function $A_0(xyzt)$ admits of such a formal extension involving but a single new formal element.*

Proof: The term scalar means that under an arbitrary rotation the functional form is invariant. [This is most evident using Dirac's functional notation: $A_0(xyzt)=\langle xyzt|A_0\rangle$. "Scalar" then means:

arbitrary rotation

$$\Rightarrow (|A_0\rangle \to |A_0'\rangle = |A_0\rangle).]$$

Hence we may use the multipole expansion theorem, exactly as discussed for $\alpha Z/|\vec{r}-\vec{a}|$, and conclude that we have but two types of multipole: $l=$even and $l=$odd. There may be (denumerably) infinitely many independent multipoles in general, but they always divide into the two classes stated. *Thus the action of $P_{\rm ext}$ on an arbitrary scalar electromagnetic field is well defined.* Q.E.D.

Similarly we may extend the analysis to the vector potential, $\vec{A}(xyzt)$, and using the vector spherical harmonic expansion[26] repeat the analysis for $A_0(xyzt)$, *mutatis mutandis*. (This whole analysis is, in effect, little more than the familiar result that functions may be split into odd and even parts.)

We have thus proven the following: *Formal adjunction of elements corresponding to the concept of external parity,* in particular, $\eta_3 \to \eta_3^{\rm ex} \equiv \eta_3 P_{\rm ext}$, *allows one to implement the relations*

$$[\eta_3^{\rm ex}, \pi_0]_- = 0\,, \qquad (51a)$$

$$[\eta_3^{\rm ex}, \vec{\sigma}\cdot\vec{\pi}]_+ = 0\,. \qquad (51b)$$

We conclude, moreover, the following proposition.

Proposition I. The operators $\mathcal{O}_\pm$ defined by

$$\mathcal{O}_\pm \equiv \eta_3^{\rm ex}(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{\pi}) \pm m \qquad (52)$$

constitute a factorization of each, and hence both, of Kramers's two equations and of the iterated Dirac equation for canonically incorporated electromagnetic fields.

Proof. The proposition is equivalent to the following result:

$$\mathcal{O}_+\mathcal{O}_- = \mathcal{O}_-\mathcal{O}_+ = \pi\cdot\pi + e\vec{\sigma}\cdot(i\rho_1\vec{E}+\vec{B}) - m^2\,, \qquad (53)$$

which is evident upon using Eqs. (51a) and (51b). Q.E.D.

(2) Part of the unease at this formal adjunction of new elements into the system stems from the fact that "external parity" might really imply adjoining a whole universe of (undefined, even undefinable) new operators. This is a valid criticism, and one really must limit the adjunction to a fixed list of elements defined at the outset, once and for all. *So far we mean only that $A(xyzt)$ has been split into two parts, with an operator $P_{\rm ext}$ defined to implement the split.*

Let us remark that after we have reanalyzed the whole structure in Sec. V – using Wigner's analysis of the Poincaré group – we can, and shall, reformulate the problem anew in a way that appears perhaps less *ad hoc* (Sec. VI).

(3) Another part of the unease stems from the feeling that the whole idea is a sort of trick "not letting the left hand know what the right hand is doing." Let us answer this unease by an example taken from a more familiar subject. Suppose we were presented with the operator $\vec{r}$ – where the symbols mean just what they normally do – and were told to answer, yes or no, the following question: Is $\vec{r}$ a vector? *We can make no answer for the question is meaningless.*

For example, we might correctly answer *yes*, meaning thereby that we *assume* that "vector" means $L=1$ under the rotation operator $\vec{L}=\vec{r}\times\vec{p}$.

But we can also answer *no*, if we assume that – as in the famous Corben-Schwinger formulation[27] – we adjoin to our rotation operator the unit spin angular momentum operator $S_i \equiv e_{ijk}$, and then define the (total) rotation operator to be $\vec{J}=\vec{L}+\vec{S}$. Under this operator, the operator $\vec{r}$ is an invariant operator having total angular momentum zero.

It is in precisely the same way as in the Corben-Schwinger example that we have adjoined formal elements that enable us to say – consistently – that under the "combined parity," $\eta_3^{\rm ex}$, A_0 is always even and $\vec{A}$ is always odd.

In this way we arrive at four equations which incorporate the interaction with an arbitrary electromagnetic field; these equations we still denote as A_2 (since they clearly belong to the second alternative in the dichotomy):

$$[\eta_3^{\rm ex}(p_0 - eA_0 - \vec{\sigma}\cdot(\vec{p}-e\vec{A})) \pm m]\psi = 0\,, \qquad (54a)$$

$$[\eta_3^{\rm ex}(p_0 - eA_0 + \vec{\sigma}\cdot(\vec{p}-e\vec{A})) \pm m]\psi = 0\,, \qquad (54b)$$

where $\eta_3^{\rm ex}=\eta_3 P_{\rm ext}$, and we have written out separately the two cases $\rho_1 \to \pm 1$.

It is useful to write out in detail, and in several ways, Eq. (54). First, write

$$A_0(\vec{x},t) = A_0^{\rm even}(\vec{x},t) + A_0^{\rm odd}(\vec{x},t),$$

with

$$A_0^{\rm even}(\vec{x},t) = \tfrac{1}{2}[A_0(\vec{x},t) + A_0(-\vec{x},t)]\,,$$

$$A_0^{\rm odd}(\vec{x},t) = \tfrac{1}{2}[A_0(\vec{x},t) - A_0(-\vec{x},t)]\,,$$

and similarly for $\vec{A}(\vec{x}, t)$. Then

$$\eta_3 A_0^{\text{even}}(\vec{x}, t)\eta_3 = A_0^{\text{even}}(\vec{x}, t),$$

$$\eta_3 A_0^{\text{odd}}(\vec{x}, t)\eta_3 = -A_0^{\text{odd}}(\vec{x}, t),$$

and similarly for $\vec{A}^{\text{odd}}$ and $\vec{A}^{\text{even}}$. Finally

$$P_{\text{ext}} A_0^{\text{even}} P_{\text{ext}} = A_0^{\text{even}},$$

$$P_{\text{ext}} A_0^{\text{odd}} P_{\text{ext}} = -A_0^{\text{odd}};$$

hence

$$\eta_3^{\text{ex}} A_0(\vec{x}, t)\eta_3^{\text{ex}} = A_0(\vec{x}, t).$$

Also

$$P_{\text{ext}} \vec{A}^{\text{even}} P_{\text{ext}} = -\vec{A}^{\text{even}},$$

$$P_{\text{ext}} \vec{A}^{\text{odd}} P_{\text{ext}} = \vec{A}^{\text{odd}};$$

hence

$$\eta_3^{\text{ex}} \vec{A}(\vec{x}, t)\eta_3^{\text{ex}} = -\vec{A}(\vec{x}, t).$$

It is clear from Eq. (54) that not only do A^0 and $\vec{A}$ possess parts both even and odd for P_{ext}, but that the same is true for $\psi(\vec{x}, t)$. We may write Eq. (54a), splitting ψ as $\psi = \psi^{(e)} + \psi^{(o)}$:

$$\eta_3^{\text{ex}} \{i\partial_0 + eA_0^{\text{odd}} + eA_0^{\text{even}} + \vec{\sigma}\cdot(i\vec{\partial} + \vec{A}^{\text{odd}} + \vec{A}^{\text{even}})\}(\psi^{(e)} + \psi^{(o)}) = m(\psi^{(e)} + \psi^{(o)}). \tag{55}$$

Collecting on both sides of (55) the terms which are even and those which are odd under P_{ext}, one obtains two equations which can be unified as one equation for the four-component object $\phi = \binom{\psi^{(e)}}{\psi^{(o)}}$:

$$\left[\begin{pmatrix}\eta_3 & 0\\ 0 & -\eta_3\end{pmatrix}\left\{i\partial_0 - e\begin{pmatrix}A_0^{\text{even}} & A_0^{\text{odd}}\\ A_0^{\text{odd}} & A_0^{\text{even}}\end{pmatrix} + \vec{\sigma}\cdot\left[i\vec{\partial} - e\begin{pmatrix}\vec{A}^{\text{odd}} & \vec{A}^{\text{even}}\\ \vec{A}^{\text{even}} & \vec{A}^{\text{odd}}\end{pmatrix}\right]\right\} - m1\right]\phi = 0. \tag{54'}$$

This equation is nothing but another representation of A_2, where η_3^{ex} is represented by

$$\begin{pmatrix}\eta_3 & 0\\ 0 & -\eta_3\end{pmatrix}.$$

The existence of this four-component representation does not of course mean that we have failed to split the two-component Klein-Gordon equation with A_2. We shall return to this alternative representation after discussing the Poincaré invariance of A_2 [cf. also Remark after Sec. VI].

B. Poincaré Invariance of A_2

The factorization of the iterated Dirac equation, using the second alternative for the Dirac dichotomy, has led us to the following equations:

$$[\eta_3^{\text{ex}}(\pi_0 - \rho_1\vec{\sigma}\cdot\vec{\pi}) \pm m]\psi(xyzt) = 0. \tag{56}$$

We have proven that this system iterates into the same result as the Dirac equation, Eq. (56). Thereby we are assured that much, if not all, of the physics of the Dirac equation carries over to A_2, *if A_2 is otherwise acceptable*. This proviso means, in effect, that the system A_2 must not only be self-consistent, it must accord with the requirement of Poincaré invariance.

This immediately poses a problem[28]: The momentum operator $\vec{p}$ does not commute with the free-field Hamiltonian:

$$H = \rho_1\vec{\sigma}\cdot\vec{p} \pm \eta_3^{\text{ex}} m, \tag{57}$$

since

$$[H, \vec{p}] = \pm 2\eta_3^{\text{ex}} m\vec{p}. \tag{58}$$

The way out of this impasse is probably evident: We must *also* adjoin as a parity-carrying operator the distance through which the system is displaced – even if infinitesimal – since displacements are necessarily *polar* vectors. To see most easily that this suffices consider a finite displacement by a distance $\vec{d}$:

$$U(\vec{d}) = e^{-i\vec{d}\cdot\vec{p}}. \tag{59}$$

One now has the desired result

$$U^{-1}HU = H, \tag{60}$$

since, by adjunction of $\vec{d}$ as a parity-carrying operator, one has

$$U^{-1}\eta_3^{\text{ex}} U = \eta_3^{\text{ex}}. \tag{61}$$

This result implies an infinitesimal form of Eq. (61) which has unusual properties, namely,

$$\text{Eq. (61)} \Rightarrow [\vec{d}\cdot\vec{p}, \eta_3^{\text{ex}}] = 0, \tag{62}$$

but

$$\vec{d}\cdot[\vec{p}, \eta^{\text{ex}}] \neq 0. \tag{62'}$$

(*Remark*. At first glance Eqs. (62) and (62′) are completely at variance with well-known group-theoretic structures wherein the *generators of a group necessarily (by definition) commute with the parameters defining finite group elements*. The flaw in this otherwise correct assertion is that *parity is not an element of the* (continuous) *translation group generated by $\vec{p}$, but is a discrete element of a distinct Z_2 group*. The distance $\vec{d}$ is indeed a parametric element for T_3 (since $[\vec{p}, \vec{d}] = 0$), but this has nothing to do with the behavior of $\vec{d}$ under P_{ext}.)

Having "rescued" displacement invariance – thereby distinguishing "displacement invariance" from "invariance under commutation by $\vec{p}$" – we can proceed to construct a Poincaré-invariant interpretation of Eq. (57).

The method to be followed is a purely algebraic one in which one first defines a (pseudo) inner product,[29] then constructs explicit operators $\vec{J}$ and $\vec{K}$ satisfying the Lorentz group commutation relations leaving this inner product invariant. If the Hamiltonian (57) is covariant under $\vec{J}$ and $\vec{K}$ and invariant under $\vec{p}$, then Eq. (57) integrates into a Poincaré group representation, completing the demonstration.

The first step is to define an appropriate pseudo inner product. We choose the form suggested already in Sec. III (with $\eta_3 \to \eta_3^{ex}$), that is,

$$\langle \overline{\psi} | \varphi \rangle \equiv \int d^3x (\overline{\psi} | \varphi), \tag{63a}$$

where

$$\overline{\psi} \equiv \psi^\dagger \eta_3^{ex} \tag{63b}$$

and $(\cdots | \cdots)$ denotes the usual spinor mapping into complex valued functions.

Let us verify some properties of this inner product. First we note that the inner product of Eq. (63) is invariant under space-time displacements.

[*Proof:*

$$\psi \to \psi' \equiv U\psi = \exp(-id\cdot p)\psi, \tag{64}$$

$$\Rightarrow \overline{\psi} \to (\overline{\psi})' = (U\psi)^\dagger \eta_3^{ex} \tag{65}$$

$$= \overline{\psi}\eta_3^{ex} U \eta_3^{ex} \tag{66}$$

$$= \overline{\psi}. \quad \text{Q.E.D.}] \tag{67}$$

The rotation operator is given by the operator $\vec{J}$, defined by

$$\vec{J} = \vec{L} + \tfrac{1}{2}\vec{\sigma}, \tag{68}$$

which obeys

$$\vec{J} \times \vec{J} = i\vec{J}, \tag{69}$$

as required.

We now verify that the inner product of Eq. (63) is invariant under spatial rotations.

[*Proof:*

$$\psi \to \psi' \equiv e^{-i\theta\hat{n}\cdot\vec{J}}\psi, \tag{70}$$

$$\Rightarrow \overline{\psi} \to (\overline{\psi})' = \overline{\psi}\eta_3^{ex} e^{+i\theta\hat{n}\cdot\vec{J}} \eta_3^{ex}, \tag{71}$$

$$= \overline{\psi}. \quad \text{Q.E.D.}] \tag{72}$$

Note that in the step from Eq. (71) to Eq. (72) we explicitly used the fact that the rotation parameters $\theta\hat{n}$ define an *axial vector*, so that $[\hat{n}, P_{ext}] = 0$.

Corresponding to the orbital angular momentum generator $\vec{L}$ we define an orbital boost operator $\vec{N}$, obeying the relations

$$[L_i, N_j] = i\epsilon_{ijk}N_k, \tag{73}$$

$$[\vec{\sigma}, \vec{N}] = 0, \tag{74}$$

$$[N_i, N_j] = -i\epsilon_{ijk}L_k, \tag{75}$$

and impose the requirement

$$\vec{N}^\dagger = \vec{N}. \tag{76}$$

(These relations are precisely those imposed in the usual Lorentz covariance discussion of the Dirac equation.)

For later use, we note here that the operator $\vec{N}$ has odd parity under the internal parity operation:

$$[\eta_3, \vec{N}]_+ = [P_{int}, \vec{N}]_+ = 0. \tag{77}$$

The critical step now is to define the spin part of the (total) boost operator $\vec{K}$. We define this to be

$$\vec{K} \equiv \vec{N} + \tfrac{1}{2}i\vec{\sigma}. \tag{78}$$

It is easily verified that $\vec{J}$ and $\vec{K}$ satisfy the defining Lorentz group commutation relations. That is

(a) $$\vec{J}\times\vec{J} = i\vec{J}, \tag{69}$$

(b) $$\begin{aligned}[J_i, K_j] &= [L_i, N_j] + i[\tfrac{1}{2}\sigma_i, \tfrac{1}{2}\sigma_j] \\ &= i\epsilon_{ijk}(N_k + \tfrac{1}{2}i\sigma_k) \\ &= i\epsilon_{ijk}K_k,\end{aligned} \tag{79}$$

and

(c) $$\begin{aligned}[K_i, K_j] &= [N_i, N_j] + [\tfrac{1}{2}i\sigma_i, \tfrac{1}{2}i\sigma_j] \\ &= -i\epsilon_{ijk}L_k - \tfrac{1}{2}i\epsilon_{ijk}\sigma_k \\ &= -i\epsilon_{ijk}J_k.\end{aligned} \tag{80}$$

One next verifies that the inner product is invariant under boosts. A finite boost is defined to be

$$B \equiv e^{i\chi\hat{v}\cdot\vec{K}}, \tag{81a}$$

$$= U(i\chi\hat{v}\cdot N)W(-\tfrac{1}{2}\chi\hat{v}\cdot\vec{\sigma}), \tag{81b}$$

$$= WU, \quad \text{using Eq. (74)}. \tag{81c}$$

It is essential at this stage to note explicitly that the boost parameters $\chi\hat{v}$ define a *polar vector* (related to the velocity vector) and hence that we must require

$$[\chi\hat{v}, \eta_3^{ex}]_+ = [\chi\hat{v}, P_{ext}]_+ = 0. \tag{82}$$

This last remark makes straightforward the verification of the invariance of the inner product under arbitrary boosts. In detail one finds

(a) invariance under $U(i\chi\hat{v}\cdot\vec{N})$:

$$\begin{aligned}\eta_3^{ex}U^\dagger\eta_3^{ex} &= \eta_3^{ex}U^{-1}\eta_3^{ex}, \quad \text{by Eqs. (76) and (81a)} \\ &= \eta_3 U \eta_3, \quad \text{by Eq. (82)} \\ &= U^{-1}, \quad \text{by Eq. (77)};\end{aligned} \tag{83a}$$

(b) invariance under W:

$$\begin{aligned}\eta_3^{ex}W^\dagger\eta_3^{ex} &= \eta_3^{ex}W\eta_3^{ex}, \quad \text{by (81a)} \\ &= \eta_3 W^{-1}\eta_3, \quad \text{by (82)} \\ &= W^{-1}, \quad \text{since } [\eta_3, \vec{\sigma}] = 0.\end{aligned} \tag{83b}$$

The inner product, Eq. (63), is accordingly invariant under the Lorentz group $\mathfrak{L}$ generated by $\vec{J}$ and $\vec{K}$. Q.E.D.

Finally we verify the invariance of Eq. (56). Clearly, since Eq. (56) is invariant for time and space translations as well as for rotations, we need only verify the invariance under boosts. Consider the operator $\mathcal{O}=\eta_3^{\text{ex}}(\pi_0-\rho_1\vec{\sigma}\cdot\vec{\pi})$. The transformation

$$g:\psi\to\psi'\equiv S(g)\psi\,, \qquad (84)$$

where S is a Lorentz boost transformation generated by $\vec{K}$ acting on ψ, induces on $\mathcal{O}$ the transformation

$$\text{boost } g:\mathcal{O}\to\mathcal{O}'\equiv S^{-1}\mathcal{O}S\,. \qquad (85)$$

The spin and space-time parts of the operator $S(g)$ commute and define transformations that may be considered separately:

$$S(g)=S_{\text{spin}}S_{\text{orb}}=S_{\text{orb}}S_{\text{spin}}\,. \qquad (86)$$

We have already verified, in Eq. (78), that for these two operators we have the relations

$$\text{(a)}\quad S_{\text{spin}}{}^{-1}\eta_3^{\text{ex}}=\eta_3^{\text{ex}}S_{\text{spin}}\,, \qquad (87\text{a})$$

$$\text{(b)}\quad S_{\text{orb}}{}^{-1}\eta_3^{\text{ex}}=\eta_3^{\text{ex}}S_{\text{orb}}{}^{-1}\,. \qquad (87\text{b})$$

Consider now the induced action of $S_{\text{orb}}(g)$ on the space-time four-vector p, or (for four-vector external electromagnetic fields) the space-time four-vector π:

By definition, we have

$$g:\pi_\mu\to\pi'_\mu\equiv S_{\text{orb}}{}^{-1}(\pi_\mu)S_{\text{orb}}=a_{\mu\nu}(g)\pi_\nu\,, \qquad (88)$$

where $a_{\mu\nu}(g)$ is the nonunitary space-time boost matrix corresponding to the boost g. (Note that the adjoined parity carrying variables in A do not effect this result at all.)

Similarly the induced action of the spin transformation operator $S_{\text{spin}}(g)$ on the spin four-vector is

$$\begin{aligned}\sigma_\mu&\equiv 1, &&\mu=0\\ &\equiv\rho_1\sigma_k, &&\mu=k=(123)\,,\end{aligned} \qquad (89)$$

$$g:\sigma_\mu\to\sigma'_\mu\equiv S_{\text{spin}}(\sigma_\mu)S_{\text{spin}},\quad \text{using (87b)} \qquad (90)$$

and hence

$$g:\sigma_\mu\to\sigma'_\mu=a_{\mu\nu}(g)\sigma_\nu\,. \qquad (91)$$

It follows that

$$g:\mathcal{O}\to\mathcal{O}'=S^{-1}\eta_3^{\text{ex}}(\pi_0-\rho_1\vec{\sigma}\cdot\vec{\pi})S=\mathcal{O}\,, \qquad (92)$$

that is, $\mathcal{O}$ is invariant under $\mathfrak{L}$, and (since $\pm m$ is obviously invariant) Eq. (56) is invariant as required. Q.E.D.

The very close similarity between this explicit proof and the usual proof of the Lorentz invariance of (A_1) is apparent; in both cases the proof goes through because in A_1, ρ_1 and ρ_3 change the various signs in the transformation operators properly, and in A_2, $\hat{v}$ and P_{ext} *change the same signs as in* A_1 *and in the same way*.

It is useful to elaborate our language and to label (with a Greek letter) the doubling of the space that has occured by adjoining the external parity variables. It so happens that there are available three ancient Greek letters,[30] stigma (to be distinguished from the terminal sigma), qappa, and sampi, which had the numerical significance of 6, 90, and 900, respectively. Since one can relate (numerologically) the known leptons to 6, 90 and 900 appearing to be hopelessly superfluous, we will use stigma, written as ς, for this new ± attribute. Stigma, appropriately enough, also denotes "mark" or "conjugate" (in mathematics). (There are other meanings of stigma that may possibly prove useful in the long run.) Stigma space will be covered by the four Hermitian Pauli type 2×2 matrices $\varsigma_0=1$, ς_i. Instead of adjoining the implicit two-dimensional parity space carried by the Poincaré group parameters $\vec{a}$ and $\vec{v}$, one can make this space *explicit*, by the adjoined operators ς_1 for $\vec{d}$, and for $\vec{v}$; and ς_3 for P_{ext}.

In terms of this new language, we can now rewrite Eq. (56) in its final form:

$$[\eta_3\varsigma_3(\pi_0-\rho_1\vec{\sigma}\cdot\vec{\pi})-m]\psi=0\,. \qquad (56')$$

Let us remark that

(a) $\eta_3\equiv P_{\text{int}}$,

with P_{int} acting *only* on the variables $\vec{x}$ and $\vec{p}$.

(b) $\varsigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$

and the Poincaré group parameters $\vec{d}$ and $\chi\hat{v}$ now explicitly become $\vec{d}\to\vec{d}\varsigma_1$ and $\chi\hat{v}\to\chi\hat{v}\varsigma_1$.

(c) The ± sign in Eq. (56) has been suppressed in precisely the same way that one chooses the "Dirac equation" to have $-m$ [in Eq. (5a)] and ignores the conjugate equation [Eq. (5b)] having opposite sign.

[The remark in (c) is not as trivial as it may appear. For the field-free case, the Dirac equation possesses both positive- and negative-energy solutions; by contrast, Eq. (56), (with $A=0$), possesses *either* positive- *or* negative-energy solutions depending on the ± sign. The sign in this latter case is therefore nontrivial, but if we unite the two cases together (i.e., introduce ς_3), an over-all ± sign now becomes as unimportant as in the Dirac equation.

Although the final result, Eq. (56′), accords fully with the results stated in Ref. 4, nevertheless the concept of stigma space is now considerably clarified, andgeneralized, as compared to the original discussion in that paper.]

The interesting thing about stigma space is that for A_2, both for even environment and without external fields, ς_3 can be taken sharp:

$$\varsigma_3\psi = \pm\psi . \tag{93}$$

To repeat: *ς_3 can be taken to be sharp in every fixed reference frame, but it "rotates" under either displacements or boosts.* It is in this precise sense that we say that stigma space splits into two-component equations in every fixed frame of reference.

As we shall see in Sec. VI, if one did not have the possibility of categorizing chirality (via weak interactions) stigma space would be abstractly isomorphic to Dirac ρ space.

It is useful to state here explicitly some further properties of Eq. (56′) which are of help in applications.

The first point is to note that – relative to the inner product of Eq. (63) – the operators

$$\mathcal{O}_\mu \equiv \eta_3\varsigma_3(1, \rho_1\vec{\sigma})$$

transform as a four-vector. [The proof has already been given in Eqs. (84)–(88).]

Using this fact, one can now show that there exist conserved currents in A_2. If we use the spinor product, $(\psi|\mathcal{O}_\mu|\psi)$, to define the four-vector current densities

$$j_\mu(x) = (\psi(x)|\mathcal{O}_\mu|\psi(x)) ,$$

then it follows from Eq. (56′) that these currents obey a conservation law[31]:

$$\partial^\mu j_\mu = 0 .$$

This result is important, since it shows that *in every fixed reference frame the two currents having sharp stigma are separately conserved.* (This follows because ς_3 is sharp in every fixed frame and hence the current j_μ splits in every frame into the two distinct stigma components. Q.E.D.)

[Unless care is expressed in regard to the various adjoined parity spaces, one can obtain seemingly paradoxical results in deriving conserved currents from Eq. (56′). Thus, for example, the operator η_3 is self-adjoint with respect to an inner product integrating over all three-space, but not with respect to a product $\psi^+\varphi$. It is essential in verifying the properties of the currents j_μ to introduce all parity spaces as explicit two-component spaces and to define the spinor product in such a way that all parity space operations are explicitly Hermitian.]

Finally let us remark that Eq. (56′) can be used to discuss the Heisenberg form of the equations of motion for A_2, in a manner that exactly parallels the usual discussion for A_1. In this way, one finds that the four-velocity is $v = (1, \rho_1\vec{\sigma})$, and that with respect to this four-velocity the Lorentz force law obtains for general external electromagnetic fields. Thus – to the extent that the one-particle approach is valid – A_1 and A_2 describe the same physics for spin-$\frac{1}{2}$ particles in external electromagnetic fields.

V. GROUP-THEORETIC ANALYSIS

The main purpose of this section is to give an explicit construction of our equation starting from an irreducible unitary representation of the covering group of the Poincaré group.[32] Such a construction once more establishes that A_2 is invariant for Poincaré transformations, and shows also that this equation describes particles with nonzero mass and spin $\frac{1}{2}$. This section consists of two parts. Firstly, we review some of the results concerning the unitary representation of the Poincaré group. Secondly, we use this material to construct the Dirac equation, and also equations A_2. Finally we show the abstract isomorphism (but not the *identity*) of our equations to the Dirac equation.

A. Unitary Representations of the Poincaré Group

We use the name Poincaré group, denoted $\mathcal{P}$, for the group of proper orthochronous inhomogeneous Lorentz transformations. Elements of the covering group of $\mathcal{P}$ are denoted by $(d;A)$, where d represents a space-time translation and A a complex 2×2 matrix of unit determinant. Wigner showed[32] that for quantum mechanics, invariance for the group $\mathcal{P}$ leads to a unitary representation of the covering group $\mathcal{P}$. The irreducible unitary representation which corresponds to a particle of spin $\frac{1}{2}$ and mass m is equivalent to[33]

$$[U(d;A)\phi]_i(p') = e^{ip'd}\sum_{j=1}^{2}\left[\left(\left(\frac{\{p'\}}{m}\right)^{1/2}\right)^{-1} A\left(\frac{\{p\}}{m}\right)^{1/2}\right]_{ij}\phi_j(p) , \tag{94}$$

where

$$p' = \Lambda(A)p, \quad p^2 = p'^2 = m^2, \quad p^0 > 0 ,$$

and

$$\frac{\{p\}}{m} = \frac{1}{m}\left\{p^0\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + p^1\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + p^2\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + p^3\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right\}. \tag{95}$$

Also useful is the relation

$$\left(\frac{\{p\}}{m}\right)^{-1} = \frac{1}{m}\left\{p^0\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - p^1\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} - p^2\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} - p^3\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right\}. \tag{96}$$

The functions $\phi_i(p)$ clearly satisfy the Klein-Gordon equation

$$(p^2 - m^2)\phi_i(p) = 0\,. \tag{97}$$

The invariant inner product is given by

$$(\phi, \psi) = \int \frac{d\vec{p}}{p^0} \sum_{i=1}^{2} \phi_i^*(p)\psi_i(p)\,. \tag{98}$$

For what follows it is useful to introduce a self-adjoint operator $\underline{\Sigma}$, defined by

$$(\underline{\Sigma}\phi)_i(p^0, \vec{p}) \equiv \phi_i(p^0, -\vec{p}), \quad \underline{\Sigma}^2 = 1\,. \tag{99}$$

The operator $\underline{\Sigma}$ commutes with the operator p_0, but anticommutes with the operators $\vec{p}$. The functions $\phi_i(p)$ are not convenient if one wishes to describe interactions[34]: The reason for this is that the transformation (94) depends on $\vec{p}$. Therefore one defines other functions

$$\phi_N(p) \equiv \left(\frac{\underline{\{p\}}}{m}\right)^{1/2} \phi(p)\,, \tag{100}$$

which transform according to

$$(U(a;A)\phi_N)_i(p') = e^{ip'a} \sum_{j=1}^{2} A_{ij}\phi_{Nj}(p)\,. \tag{101}$$

The invariant inner product (ϕ, ψ) can now be written as

$$\int \frac{d^3p}{p^0}\, \phi_N^\dagger(p)\left(\frac{\underline{\{p\}}}{m}\right)^{-1} \psi_N(p)\,. \tag{102}$$

For completeness it is useful to point out that the functions

$$\phi_M(p) = \left(\frac{\underline{\{p\}}}{m}\right)^{-1} \phi(p) \tag{103}$$

transform according to

$$[U(d,A)\phi_M]_i(p') = e^{ipd} \sum_{j=1}^{2} \left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} A^* \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]_{ij} \phi_{Mj}(p)\,. \tag{104}$$

B. Construction of the Dirac Equation and of Equation A_2

A solution to the Dirac equation can be obtained by constructing[35] a four-component wave function $\phi_D(p)$ from $\phi_N(p)$:

$$\phi_D(p) = \begin{pmatrix} \phi_N(p) \\ \phi_M(p) \end{pmatrix} = \begin{pmatrix} \phi_N(p) \\ (\{p\}/m)^{-1}\phi_N(p) \end{pmatrix}. \tag{105}$$

The four-component function $\phi_D(p)$ satisfies

$$\begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & -\dfrac{\underline{\{p\}}}{m} \\ -\left(\dfrac{\underline{\{p\}}}{m}\right)^{-1} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \phi_D(p) = 0\,, \tag{106}$$

which can be rewritten as

$$(\gamma^\mu p^\mu - m)\phi_D(p) = 0\,,$$

$$\gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \tag{107}$$

which is a somewhat unfamiliar representation of the Dirac equation. In terms of $\phi_D(p)$ the inner product (98) reads

$$\frac{1}{2}\int \frac{d\vec{p}}{p^0}\, \phi_D^\dagger(p)\gamma^0\psi_D(p) = \frac{1}{2}\int \frac{d\vec{p}}{p^0}\, \phi_D^\dagger(p)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\psi_D(p)\,. \tag{108}$$

The function $\phi_D(x)$ is defined by

$$\phi_D(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3p}{p^0}\, e^{-ipx}\, \phi_D(p)\,. \tag{109}$$

This is a positive-energy solution to the Dirac equation. Negative-energy solutions can be included, by a procedure analogous to the one carried out here, starting, instead of with (100), with

$$\phi_N(p) = \left(\frac{\underline{\{p\}}}{m}\right)^{1/2} \phi(p) + \left(\frac{\underline{\{p\}}}{m}\right)^{1/2} \phi(-p),$$

$$p^0 > 0, \quad -p^0 < 0\,. \tag{110}$$

The transformation properties of $\phi_D(x)$ and $\phi_D(p)$ follow from those of $\phi(p)$.

Equation A_2 can be constructed in a similar way. Consider

$$\phi_\varsigma(p) = \begin{pmatrix} \phi_N(p) \\ \phi_L(p) \end{pmatrix} = \begin{pmatrix} \phi_N(p) \\ (\{p\}/m)\underline{\Sigma}\phi_N(p) \end{pmatrix} \tag{111}$$

and alternatively

$$\phi_\varsigma(p) = \begin{pmatrix} \phi_M(p) \\ \phi_K(p) \end{pmatrix} = \begin{pmatrix} \phi_M(p) \\ (\{p\}/m)^{-1}\underline{\Sigma}\phi_M(p) \end{pmatrix}. \tag{112}$$

The functions $\phi_\varsigma(p)$ satisfy equations (A_2). First, remark that

$$\frac{\underline{\{p\}}}{m}\underline{\Sigma}\phi_L = \frac{\underline{\{p\}}}{m}\underline{\Sigma}\,\frac{\underline{\{p\}}}{m}\underline{\Sigma}\phi_N = \frac{\underline{\{p\}}}{m}\left(\frac{\underline{\{p\}}}{m}\right)^{-1}\underline{\Sigma}^2\phi_N = \phi_N\,.$$

Hence

$$\begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & -\dfrac{\underline{\{p\}}}{m}\underline{\Sigma} \\ -\dfrac{\underline{\{p\}}}{m}\underline{\Sigma} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \phi_\varsigma(p) = 0\,. \tag{113}$$

In terms of the functions ϕ_ς the inner product (98) reads

$$\frac{1}{2}\int \frac{d\vec{p}}{p^0}\, \phi_\varsigma^\dagger(p)\begin{pmatrix} 0 & \Sigma \\ \underline{\Sigma} & 0 \end{pmatrix}\psi_\varsigma(p)\,. \tag{114}$$

Note that $\phi_\varsigma(p)$ can be expressed as

$$\left\{\begin{pmatrix} 1 & 0 \\ 0 & \underline{\Sigma} \end{pmatrix}\phi_D\right\}(p) = \phi_\varsigma(p)\,. \tag{115}$$

The transformation properties of ϕ_ς follow, just as those for ψ_D, from those of ϕ. For space-time translations, for instance, we find

$$\phi_\varsigma(p)=\begin{pmatrix}\phi_N(p)\\ \phi_L(p)\end{pmatrix}\xrightarrow[(\vec{a},a_0)]{} e^{ip^0a^0}\begin{pmatrix}e^{-i\vec{p}\cdot\vec{a}}\phi_N(p)\\ e^{+i\vec{p}\cdot\vec{a}}\phi_L(p)\end{pmatrix}. \qquad (116)$$

Next we define

$$\phi_\varsigma(x)=\frac{1}{(2\pi)^{3/2}}\int\frac{dp}{p^0}\,e^{-ipx}\,\phi_\varsigma(p)\,, \qquad (117)$$

and introduce η_3. Equation (113) can then be rewritten in x space as

$$\left[m\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}+\begin{pmatrix}0 & i\left\{\frac{\partial}{\partial t}\begin{pmatrix}1&0\\0&1\end{pmatrix}-i\vec{\partial}\cdot\vec{\sigma}\right\}\eta_3\\ i\left\{\frac{\partial}{\partial t}\begin{pmatrix}1&0\\0&1\end{pmatrix}-i\vec{\partial}\cdot\vec{\sigma}\right\}\eta_3 & 0\end{pmatrix}\right]\psi_\varsigma(x)=0\,, \qquad (118)$$

which we shall abbreviate as

$$\left[m1+\left(i\frac{\partial}{\partial t}-i\vec{\partial}\cdot\vec{\sigma}\right)\begin{pmatrix}0&\eta_3\\ \eta_3&0\end{pmatrix}\right]\psi_\varsigma(x)=0\,. \qquad (119)$$

This equation can be rewritten as

$$\left[m1+\begin{pmatrix}0&\eta_3\\ \eta_3&0\end{pmatrix}\left(i\frac{\partial}{\partial t}+i\vec{\partial}\cdot\vec{\sigma}\right)\right]\psi_\varsigma(x)=0\,. \qquad (120)$$

[Equation (120) is identical to (54b); similarly from (112) one can obtain (54a) or (54′).] Equation (120) possesses a local conserved current which suggests coupling to the electromagnetic field by

$$\left[m1+\begin{pmatrix}0&\eta_3\\ \eta_3&0\end{pmatrix}\left\{i\frac{\partial}{\partial t}-e\begin{pmatrix}A_0(\vec{x},t)&0\\0&A_0(-\vec{x},t)\end{pmatrix}+\vec{\sigma}\cdot\left(i\vec{\partial}-e\begin{pmatrix}\vec{A}(\vec{x},t)&0\\0&-\vec{A}(-\vec{x},t)\end{pmatrix}\right)\right\}\right]\psi_\varsigma(\vec{x},t)=0\,. \qquad (121)$$

This equation can be changed into A_2 by a rotation in stigma space.

It follows then that one may derive either A_1 or A_2 directly from the Wigner's analysis of the Poincaré group. Such an analysis leads to A_2 with stigma space fully explicit.

It is worthwhile to notice that with Eq. (115) we find

$$\left\{\begin{pmatrix}1&0\\0&\eta_3\end{pmatrix}\phi_D\right\}(x)=\phi_\varsigma(x)\,. \qquad (122)$$

This suggests a one-to-one mapping of our equation into the Dirac equation. This mapping is the subject of our next section. The mapping applies also to the interaction with a general electromagnetic field ensuring that the results of electrodynamics will be the same for A_2 as for Dirac's equation.

VI. AN INVERTIBLE MAPPING OF A_1 ONTO A_2

In this section we demonstrate that it is possible to find an invertible mapping[36] of $A_1 \to A_2$ that is both an into and an onto map. This mapping demonstrates the abstract isomorphism of stigma space with Dirac ρ space, so far as external electromagnetic interactions are concerned. After establishing this mapping we discuss that nontheless there is a difference between A_1 and A_2 and that one should not assert that A_1 and A_2 are "equivalent." At the end of this section we briefly discuss two other, unsuccessful, attempts to obtain a Poincaré-invariant two-component equation for massive spin-$\frac{1}{2}$ particles.

First we verify the existence of the one-to-one map. We shall do this only for the equation without electromagnetic fields. (The derivation with electromagnetic field is identical, except that the equations become longer.[37])

Let us be explicit. Consider the following transformation [note that it is defined (for free fields) completely within the Dirac structure, A_1]:

$$\begin{aligned}S&\equiv\left(\frac{1+\rho_1}{2}\right)+\left(\frac{1-\rho_1}{2}\right)\eta_3\\ &=P_+ +\eta_3P_-\,,\end{aligned} \qquad (123)$$

where we have the properties

$$\begin{aligned}&(P_\pm)^2=P_\pm\,,\quad P_+ +P_-=1\,,\\ &[P_\pm,\eta_3]=0\,.\end{aligned} \qquad (124)$$

Note that $S^{-1}=S$.

Consider the Dirac equation, written in covariant form:

$$\begin{aligned}\mathcal{O}_{\text{Dirac}}\,\psi&\equiv(\rho_3p_0-i\rho_2\vec{\sigma}\cdot\vec{p}-m)\psi\\ &=0\,.\end{aligned} \qquad (125)$$

Transforming this equation by S acting on the left-hand side yields

$$S\mathcal{O}_{\text{Dirac}}S^{-1}S\psi \equiv (\mathcal{O}_{\text{Dirac}})'(\psi)' = 0\,, \tag{126}$$

where

$$S: \mathcal{O}_{\text{Dirac}} \to (\mathcal{O}_{\text{Dirac}})' \equiv S(\rho_3 p_0 - i\rho_2 \vec{\sigma}\cdot\vec{p} - m)S\,. \tag{127}$$

To evaluate the transformed operator in Eq. (127) it is helpful to divide the work into two parts.

(a) The $(\rho_3 p_0 - m)$ part: One easily finds that

$$S(\rho_3 p_0 - m)S = \eta_3 \rho_3 p_0 - m\,. \tag{128}$$

(b) The $(-i\rho_2\vec{\sigma}\cdot\vec{p})$ part: This part transforms very curiously:

$$\begin{aligned} S(-i\rho_2\vec{\sigma}\cdot\vec{p})S &= (-i\rho_2)(P_- + \eta_3 P_+)(\vec{\sigma}\cdot\vec{p})S\,, \\ &\quad \text{using Eq. (123) for } S \\ &= (-i\rho_2)(P_- + \eta_3 P_+)(P_+ - \eta_3 P_-)(\vec{\sigma}\cdot\vec{p})\,, \\ &\quad \text{using } [\eta_3, \vec{\sigma}\cdot\vec{p}]_+ = 0 \\ &= (-i\rho_2)(\eta_3\rho_1)(\vec{\sigma}\cdot\vec{p})\,, \\ &\quad \text{using Eq. (124)} \\ &= -\eta_3\rho_3\vec{\sigma}\cdot\vec{p}\,. \end{aligned} \tag{129}$$

Hence we conclude that the desired transformation on $\mathcal{O}_{\text{Dirac}}$ is

$$S: \mathcal{O}_{\text{Dirac}} \to (\mathcal{O}_{\text{Dirac}})' = \eta_3\rho_3(p_0 - \vec{\sigma}\cdot\vec{p}) - m\,. \tag{130}$$

Note now that this transformed operator only involves a *single Dirac-space operator,* ρ_3. Accordingly we may split Eq. (130) into two two-component equations by taking $\rho_3 \to \pm 1$. This miracle occurs only because of the curious transformation properties detailed in Eq. (129).

Since we do not want to conclude prematurely that Dirac space (ρ space) is *identical* to stigma space (we will justify this at the end of this section), we choose to distinguish the transformed system of Eq. (130), by relabeling the (diagonal) operator ρ_3 as ς_3. Accordingly Eq. (130) becomes

$$[\eta_3\varsigma_3(p_0 - \vec{\sigma}\cdot\vec{p}) - m]\psi = 0\,, \tag{131}$$

which is now seen to be identical with the (free-field) equation for A_2, Eq. (56′).

Thus we have established our claim that there exists an invertible mapping $S: A_1 \to A_2$. The explicit form of the mapping S demonstrates, however, much more: It shows, for example, *that Dirac space for* A_1 *is mapped onto a two-component space of external parity for* A_2. It follows that there is no physical sense in asserting that "A_1 and A_2 are equivalent" – for Dirac space is certainly not a parity space. Only in the precise abstract sense in which all two-dimensional Hilbert spaces are isomorphic can one assert that – on the basis of the results just proven – A_1 and A_2 are "equivalent."

Let us consider the transformation S from the standpoint of symmetry transformations defined on the Dirac structure. A symmetry of a physical structure is defined[38] to be a mapping of the system into itself which *either* preserves all probability amplitudes *or* carries all probability amplitudes into their complex conjugates, for each coherent subspace separately. We require furthermore that a "symmetry of the Dirac structure" leave $(\bar{\psi}|\psi)$ invariant. Poincaré transformations, in particular, are symmetries of the Dirac structure under this definition.

By contrast, *the transformation* S *is not a symmetry of the Dirac structure.*

[*Proof:* The transformation S does not leave the inner product based on $\bar{\psi} \equiv \psi^\dagger\rho_3$ invariant; in fact, one sees that

$$S: \rho_3 \to S\rho_3 S = \eta_3\rho_3\,. \quad \text{Q.E.D.}]$$

This result demonstrates moreover that the transformation S carries the inner product for A_1 into the inner product for A_2 [defined in Eq. (63)], and is accordingly a mapping: $S: A_1 \to A_2$. Because this mapping preserves commutators, we can conclude that the Poincaré representations carried by A_1 and A_2 are isomorphic.[39]

It is useful at this point to examine the effect of the mapping S on various operators. We find the following "dictionary":

A_1		A_2	
p_0	$\to$	p_0	(132)
$\vec{p}$	$\to$	$\varsigma_1\vec{p}$	(133)
$\vec{J}$	$\to$	$\vec{J}$	(134)
$K_{\text{Dirac}} = N + \frac{1}{2}i\rho_1\sigma$	$\to$	$\varsigma_1\vec{K}_{A_2}$.	(135)

[Mappings (132) and (134) are obvious from the definition of S, but the remaining two mappings require comment. To establish (133), we note that η_3 has the significance of P_{int} and hence anticommutes with p. Thus one finds

$$\begin{aligned} \vec{p} \to \vec{p}' = S\vec{p}S &= \vec{p}(P_+ - \eta_3 P_-)(P_+ + \eta_3 P) \\ &= \varsigma_1\vec{p} \end{aligned} \tag{136}$$

(upon redefining ρ_1 as ς_1). Similarly the operator $\vec{N}$ anticommutes with η_3 (since $\vec{N}$ has odd parity) and hence $\vec{N} \to S\vec{N}S = \varsigma_1\vec{N}$. Thus one finds (identifying ρ_1 as ς_1) that (135) is correct.]

This dictionary establishes in complete detail the validity of the language used in Sec. IV. In particular, one sees that the (Dirac) parameters $\vec{d}$ (spatial displacements) and $\chi\hat{v}$ (boosts) under the mapping S now acquire the operator ς_1, explicitly verifying the more intuitive arguments of Sec. IV, which concluded that, for A_2, the Poincaré parameters "$\vec{d}$" ($=\vec{d}\varsigma_1$) and "$\chi\hat{v}$" ($=\chi\hat{v}\varsigma_1$) became parity-changing operators with respect to η_3^{ex} ($=\eta_3\varsigma_3$).

It is interesting to add one more operator to our dictionary. We have already seen that the (A_1) operator ρ_3 maps under S into the (A_2) operator $\varsigma_3\eta_3$.

This establishes an interesting result: *The parity operator for the Dirac equation* ($\eta_3\rho_3$) *maps under S into the* (A_2) *operator* ς_3.

This remark verifies the statement by Good[25] to the same effect. (It is essential, to remark however, that if ρ_1 has a fixed significance, then ς_3 and Dirac parity are *not* transforms under S. See discussion below.)

We must now justify, in more detail, why we chose to distinguish stigma space from Dirac ρ space – which the mapping S seems to imply are the same. Our reasoning is based on physical considerations.

Our original aim was to factorize the Kramers equation, which presupposes that ρ_1 (chirality) is sharp. Thus the factorized system A_2 can be extended to sharp chirality. Clearly the original Dirac structure – for which chirality is not a constant of the motion – can in no sense be equivalent to the A_2 structure, which has chirality a constant of the motion. Expressed in different words, we have identified the isomorphic "Dirac space" defined by the factorization with a specific physical space (parity), such that the original Dirac space (defining chirality) is incorporated in both A_1 (four-dimensional) and in A_2 (eight-dimensional, splitting over parity and chirality into four two-dimensional subspaces).

Let us now discuss briefly two other attempts at either circumventing the requirements underlying Dirac's factorization or toward obtaining two-component equations.

Consider first the so-called "square-root" Hamiltonian, which factorizes Eq. (1) in the form

$$[p_0-(\vec{p}^2+m^2)^{1/2}][p_0+(\vec{p}^2+m^2)^{1/2}]\psi=0 \qquad (137a)$$

or

$$p_0=\pm(\vec{p}^2+m^2)^{1/2}\psi . \qquad (137b)$$

This Hamiltonian clearly violates Dirac's requirement that p appear linearly in all four components – but, of course, such a requirement has no valid *a priori* basis. Much more fundamental is the result stated by Dirac and proven in greater generality by Sucher[40]: *The square-root equation cannot be extended to include interactions without losing Lorentz invariance.* This effectively eliminates (137b) as a physically meaningful wave equation.[41]

Now consider the Pryce-Foldy-Wouthuysen (PFW) transformation.[42] Applied to the free Dirac equation this transformation leads to the result

$$H_{\text{Dirac}} \xrightarrow[\text{PFW}]{} H'\psi=\rho_3(p^2+m^2)^{1/2}\psi . \qquad (138)$$

This is clearly just a doubling in Dirac space (ρ space) of the square-root Hamiltonian.

In the light of the Dirac-Sucher result one must conclude that – despite the very beautiful analyses and deep physical insights afforded by the PFW transformation – it is unacceptable as a two-component splitting of Dirac's equation. This conclusion is borne out by the well-known difficulty in defining a PFW transformation on the Dirac equation in interaction with general electromagnetic fields. The Coulomb field problem, in particular, does not admit of a solvable PFW transformation (but rather of a series of unlimitedly many approximations).[43]

By contrast, A_2 does admit two-component solutions for arbitrary external fields; in particular, the Coulomb field problem admits of such a two-component solution, as discussed in Sec. III.

Remark. Before concluding this section it might be useful to discuss briefly a question that must surely have arisen: Why do we call A_2 a two-component system? If one recalls the theorem which asserts that every faithful unitary representation of a connected noncompact semisimple group (such as $\mathcal{P}$) is infinite-dimensional, one sees that even to call the Dirac solutions four-dimensional must involve some (unstated) convention. The convention is that one chooses to ignore dimensionality based on space-time [that is, $\psi(\vec{p})$ is infinite-dimensional in the space of $\vec{p}$, for example].

It is in precisely this same sense that we call A_2 two-dimensional (in spin space), for we have chosen to ignore the dimensionality associated with the space-time structure carried by parity. This is not an unusual convention; a scalar function $f(\vec{x})$, for example, may be split into two parts, f^{even} vs f^{odd}, but one may – or may not – choose to call $f(\vec{x})$ "one-dimensional."

Thus we may choose to designate A_2 as two-dimensional, meaning thereby that all space-time structures (p, stigma, chirality) are both sharp and ignorable. There is a sense in which these dimensional considerations are not simply a matter of convention. For example, in neither A_1 nor A_2 can spin space be taken sharp. The Dirac postulate for quantum mechanics (the existence of a complete set of commuting observables) has the implication that all states can be split by sharp variables into one-dimensional subspaces. Spin space is subsumed under J, J_z which split into one-dimensional subspaces. For Dirac space there is a genuine problem here. This space is related via

C (=conjugation) to charge. But conjugation for a one-particle theory is nonlinear (antiunitary). It is necessary to invoke second quantization to obtain a linear unitary operator and a splitting quantum number.

Thus we see that, conventional or not, the question of dimensionality does relate to genuine physical problems.

VII. CONCLUDING REMARKS

The preceding sections have all been directed toward proving that A_2 exists and is a genuine alternative to A_1. Once this has been done successfully, then really interesting questions may be considered: For example, does nature choose to make any use of A_2? Or again: Which alternative does the physical electron choose? Because the Kramers equation and $g=2$ are related, it seems clear that A_2, if it applies to anything, might apply to leptons.

In this connection it should be noted that either of the Kramers equations suffices (by the mapping of S of Sec. VI) to define a Dirac-like electrodynamics (at the one-particle level). Yet the path to the Kramers equation, by way of the iterated Dirac equation, clearly introduced a doubling via ρ_1 (chirality).

This strongly suggests that the weak interactions might play a role in utilizing the freedom afforded by A_2. Parity has all along been a very puzzling variable; even more so as its role in A_2 is difficult to grasp fully. Let us note in particular that the stigma variable ($\varsigma_3 \to \pm$), is quite peculiar in that it seems to imply a *kinematic* relationship to the concept of a frame of reference. (For example, a weak current of sharp stigma could conceivably interact in a frame fixed by hadrons.)

We mention these speculations only to motivate the belief that A_2 may have some interesting applications in physics.

There are simpler questions, which we intend to answer in the near future, namely: the role of $g=2$ in factorizing Kramers equation; field quantization; the reflection operations; and the CPT theorem in the context of A_2.

ACKNOWLEDGMENTS

This work has depended heavily on discussions with our colleagues. We would particularly like to acknowledge the help of Professor V. Telegdi, who played a decisive role in the development of the work reported here. We would also like to thank Dr. L. Staunton for the benefit of his careful critique and explanation of the concept of external parity, and Professor Bernadino Bosco for an enlightening discussion on *Paritätszitterbewegung* in A_2.

---

*Work supported in part by the National Science Foundation and the Army Research Office (Durham, N. C.).

[1]P. A. M. Dirac, Proc. Roy. Soc. (London) A117, 610 (1928); A118, 341 (1928).

[2]O. Klein, Z. Physik 37, 895 (1926); V. Fock, *ibid.* 38, 242 (1926); 39, 226 (1926); W. Gordon, *ibid.* 40, 117 (1926); J. Kudar, Ann. Physik 81, 632 (1926). It is well known that de Broglie as well as Schrödinger gave this same equation even earlier.

[3]L. C. Biedenharn and M. Y. Han, Phys. Letters 36B, 475 (1971).

[4]L. C. Biedenharn, M. Y. Han, and H. van Dam, Phys. Rev. Letters 27, 1167 (1971).

[5]L. C. Biedenharn, M. Y. Han, and H. van Dam, Lett. Nuovo Cimento 2, 730 (1971).

[6]This algebra has rather ancient origins stemming originally from the work of Sommerfeld and of Dirac. A detailed presentation (including references to earlier work) is in L. C. Biedenharn and P. J. Brussaard, *Coulomb Excitation* (Clarendon, Oxford, 1965), Chap. III.

[7]The quoted material is from Ref. 1, p. 613. The italics are not in the original, but have been added here for emphasis.

[8]C. G. Darwin, Proc. Roy. Soc. (London) 118, 654 (1928).

[9]M. Berrondo and H. V. MacIntosh, J. Math. Phys. 11, 125 (1970).

[10]L. C. Biedenharn, Phys. Rev. 126, 845 (1962).

[11]L. C. Biedenharn and N. V. V. J. Swamy, Phys. Rev. 133, B1353 (1964).

[12]It is useful to note that for $\eta = 0$, we have the relation $F_{l,\eta=0}(kr) = j_l(kr)$. Moreover, in the limit $c \to \infty$, the $F_{l,\eta}$ become just $(kr)^{-1}$ times the usual nonrelativistic Coulomb functions as discussed by Breit and collaborators.

[13]This alternation between the two Hamiltonians is a direct result of using $\eta_3 = P_{\text{int}}$. Since j is a good quantum number we can, by putting $\eta_3' = f(j)\, P_{\text{int}}$ [with $f(j) = \pm$ depending on j], *arbitrarily* divide the nondegenerate levels between the two Hamiltonians. Our original choice was $\eta_3' = S(\mathcal{K})$, which put all nondegenerate levels in one Hamiltonian; for the constant magnetic field problem this choice is not admissible and one must use $\eta_3 = P_{\text{int}}$.

[14]G. Temple, *The General Principles of Quantum Theory*, (Methuen, London, 1934), 4th ed., 1948, cf. p. 93.

[15]P. C. Martin and R. O. Glauber, Phys. Rev. 109, 1307 (1958); Egil Hylleraas, Z. Physik 164, 493 (1961); L. C. Biedenharn, Phys. Rev. 126, 845 (1962).

[16]A. Sommerfeld, *Atombau und Spektrallinien, Wellenmechanischer Ergänzungs band* (Vieweg und Sohn, Braunschweig, 1929), pp. 132-139.

[17]This question was posed by Professor Dirac in connection with invited paper BB-2 presented at the Southeastern Section of the A.P.S., Columbia, S. C., 1971 (unpublished).

[18]R. P. Feynman, Rev. Mod. Phys. 20, 367 (1948); Phys. Rev. 84, 108 (1951). See also Lecture Notes on Quantum Electrodynamics, Cornell University, 1949 (unpublished).

[19]R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193 (1958). See also the very similar work of R. E. Marshak and E. C. G. Sudarshan, *ibid.* 109, 1860 (1958); and S. S. Gershtein and Ia. B. Zel'dovich, Zh. Eksp. Teor. Fiz. 29, 698 (1955) [Soviet Phys. JETP 2, 576 (1956)].

[20]B. L. van der Waerden, cited (without explicit referencing) in Ref. 23.

[21]H. A. Kramers, *Quantum Mechanics* (Interscience, New York, 1957); see also, *Zeeman Commemorative Volume,* (Nyhoff, The Hague, 1935), p. 403.

[22]This statement is based on the following definitions for the discrete operations in the ρ_1-projected two-component spin space: $C \equiv i\sigma_2 K_0$, where $K_0 \equiv$ complex conjugation, and $P \equiv P_{\text{int}} = \eta_3$. In verifying that CP is a symmetry one explicitly assumes that $[CP, A_0]_+ = [CP, \vec{A}]_- = 0$.

[23]L. M. Brown, in *Lectures in Theoretical Physics,* edited by W. E. Brittin *et al.* (Interscience, New York, 1962), Vol. IV.

[24]L. M. Brown, Phys. Rev. 111, 957 (1958); M. Tonin, Nuovo Cimento 14, 1108 (1959). Reference 19 asserted the existence of these rules earlier.

[25]R. H. Good (private communication). See also R. H. Good, Phys. Rev. D 5, 1538 (1972); and L. C. Biedenharn, M. Y. Han, and H. van Dam, *ibid.* 5, 1539 (1972).

[26]L. C. Biedenharn and M. E. Rose, Rev. Mod. Phys. 25, 729 (1953). A more extended treatment is given in Proceedings of the NATO Summer School, Cagliari, Sardinia, Italy, 1970, edited by B. Bosco (Gordon and Breach, London, to be published). The treatment in both these references is based on the work of Ref. 27.

[27]H. C. Corben and Julian Schwinger, Phys. Rev. 58, 953 (1940).

[28]We wish to thank Professor E. Merzbacher for discussions on this point.

[29]It is essential to be explicit about this inner product. We mean by this inner product [given in Eq. (63)] first the spinorial product $(\bar{\psi}|\varphi)$ mapping two spinor functions, say $\psi(x)$ and $\phi(x)$, into a complex valued function of space-time, and secondly, an integration over all space. The first part is necessary to guarantee Hermiticity of the spin operators; the second part is necessary to achieve Hermiticity for the operators p. It is a peculiarity of the Dirac equation that this particular inner product is *indefinite* (and hence a pseudo inner product), a difficulty cleared up ultimately by "second" quantization. Let us note that a quite distinct, and customary, inner product (the *space integral of the charge density)* is used in Sec. V, and constitutes a true inner product.

[30]O. Neugenbauer, *The Exact Sciences of Antiquity* (Princeton Univ. Press, Princeton, N. J., 1952), pp. 10 and 24.

[31]It follows from this result that the (pseudo) inner product in the sense of Eq. (63) leads – for free fields – to a positive definite inner product in the sense of Sec. V, Eq. (108).

[32]E. Wigner, Ann. Math. 40, 149 (1939).

[33]A. S. Wightman, in *Dispersion Relations and Elementary Particles, les Houches Lectures, 1960,* edited by C. DeWitt and R. Omnès (Wiley, New York, 1961), p. 188.

[34]S. Weinberg, Phys. Rev. 133, B1318 (1964).

[35]M. Froissart and R. Omnes, in *High Energy Physics, les Houches Lectures, 1965,* edited by C. DeWitt and M. Jacob (Gordon and Breach, New York, 1965), p. 140.

[36]The existence of such a mapping was suggested to us by T. W. Lake (private communication).

[37]For completeness, let us remark that the transformation S, given in Eq. (123), is not the only possibility. We can choose to distinguish the parity carried by the Poincaré group parameters (implemented by adjunction of stigma space) and the parity carried by the external electromagnetic fields (implemented by adjunction of P_{ext} defined to act only on the parameters in the *EM* fields and hence also on these same parameters where they occur in the wave functions). This latter viewpoint leads to

$$S^{\text{ex}} = P_+ + \eta_3^{\text{ex}} P_- ,$$

$$\eta_3^{\text{ex}} \equiv \eta_3 P_{\text{ext}} .$$

It is readily verified that this implies the equation

$$[\eta_3^{\text{ex}} \varsigma_3 (\pi_0 - \rho_1 \vec{\sigma} \cdot \vec{\pi}) - m]\, \psi = 0 .$$

The transformation S^{ex} now implies that the Dirac parity operator $P_{\text{Dirac}} = \rho_3 \eta_3$ obeys

$$S^{\text{ex}} : P_{\text{Dirac}} \to \varsigma_3 P^{\text{ex}} ,$$

and hence (in contrast to S), Dirac parity transforms into ς_3 only in even environments (see also Ref. 25). We have not completely explored the implications of S^{ex}. We prefer to base our arguments on S [Eq. (123)], *since* the Poincaré group analysis of Sec. V leads directly to this transformation. The consistency and acceptability of S is thereby assured on firm grounds.

[38]V. Bargmann, J. Math. Phys. 5, 862 (1964); E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959), p. 251; R. M. F. Houtappel, H. van Dam, and E. P. Wigner, Rev. Mod. Phys. 37, 595 (1965).

[39]Note, however, that the mapping S preserves the invariant inner product of Sec. V, Eq. (108).

[40]J. Sucher, J. Math. Phys. 4, 17 (1963).

[41]See also M. Gell-Mann, lectures at Columbia, 1954, and Caltech, 1958 (unpublished).

[42]M. H. L. Pryce, Proc. Roy. Soc. (London) A195, 62 (1948); L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78, 29 (1948). See also S. Tani, Soryushiron kenkyu 1, 15 (1949).

[43]E. de Vries, Fortschr. Physik 18, 150 (1970).