Zitterbewegung as a subdynamics

1972 J. Phys. A: Gen. Phys. 5 1438

(http://iopscience.iop.org/0022-3689/5/10/008)

J. Phys. A: Gen. Phys., Vol. 5, October 1972. Printed in Great Britain.

Zitterbewegung as a subdynamics

J RAE

Faculté des Sciences, Université Libre de Bruxelles, Ave Roosevelt, 1050 Bruxelles, Belgium

MS received 22 May 1972

Abstract. The superoperator theory of nonequilibrium statistical mechanics is applied to the Dirac and Klein–Gordon equations. The transformation operators of this theory are found to be closely related to the Foldy–Wouthuysen transformation while the associated subdynamics corresponds to the separation into smooth motion and *zitterbewegung*.

1. Introduction

In recent years Prigogine and co-workers have developed a theory of nonequilibrium statistical mechanics, referred to below as the superoperator theory, which has as its object the derivation and transformation of kinetic equations (Prigogine *et al* 1970a, b, George 1967). The two main aspects of this work are a separation of the original time evolution into independent subdynamics and a transformation of the resulting equations into what is often called the physical particle representation. More recently these ideas have found applications outside their original domain and the present article is concerned with one such case, an application to the Foldy–Wouthuysen (FW) transformation. It is found that the latter is an example of the transformations referred to above and, furthermore, that there is a close connection between the subdynamics and the *zitterbewegung* phenomenon of the Dirac and Klein–Gordon equations.

In the sections which follow we introduce first some notions and notation from the free-particle Dirac equation and the superoperator theory. Sections 4 and 5 calculate the superoperators associated with the FW transformation and apply them to produce the transformed Hamiltonian and the subdynamics. The projector Π which gives the subdynamics turns out to be exactly the 'observable projection' of Pryce (1948). Section 6 sketches the equivalent results for the Klein–Gordon equation and the situation when an electromagnetic field is present. For the discussion of free particles we use a recent form of the superoperator theory (Rae 1972a, b) which allows exploitation of the known explicit form of the FW transformation in that case.

An outline of the direct approach is given in an Appendix for those readers familiar with the more standard form of the theory. In the case where a field is present one generally has to employ perturbation methods and there is then little to choose between the two forms of the superoperator theory.

The results of this work can be regarded as providing a somewhat unorthodox view of the Dirac and Klein–Gordon equations or, better perhaps, a rather unexpected illustration of the superoperator theory.

2. The Dirac equation

The Dirac Hamiltonian for a free electron can be written (Messiah 1966, Bjorken and Drell 1964)

$$H = m\beta + \boldsymbol{\alpha} \cdot \boldsymbol{p}. \tag{1}$$

This operator acts on 4-spinors and β, $\boldsymbol{\alpha}$ are 4×4 matrices

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \tag{2}$$

where $\boldsymbol{\sigma}$ are the usual Pauli matrices and the 1 in β is a 2×2 unit matrix. The matrices β, $\boldsymbol{\alpha}$ satisfy the anticommutation relations

$$\beta\alpha_i + \alpha_i\beta = 0 \qquad \alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{ij} \qquad \alpha_i^2 = \beta^2 = 1. \tag{3}$$

The adjoint operation here is the usual transpose of the complex conjugate so that H is self-adjoint.

For some purposes it is useful to eliminate the odd operators (ie the $\boldsymbol{\alpha}$ terms) from the Hamiltonian (1), these being the terms which couple the 'large' and 'small' components of the spinors. This can be done by a unitary transformation, the FW transformation, which in the present case takes the form (Messiah 1966)

$$U = \left(\frac{m+E}{2E}\right)^{1/2} + \beta\frac{\boldsymbol{\alpha} \cdot \boldsymbol{p}}{\sqrt{2E(E+m)}} \tag{4}$$

where E is the operator $\sqrt{(m^2 + \boldsymbol{p}^2)}$. In the FW representation the Hamiltonian (1) becomes

$$H_{\mathrm{F}} = UHU^{\dagger} = \beta\sqrt{(m^2 + \boldsymbol{p}^2)}. \tag{5}$$

The operators (4) and (5) will appear again below.

The other aspect of the Dirac Hamiltonian which we wish to mention here is the appearance of the rapidly oscillating motion, the *zitterbewegung*. It is found (Messiah 1966, Bjorken and Drell 1964, Feshbach and Villars 1958) that if a wave packet is formed by superposing positive and negative energy solutions of the Dirac equation, its motion under the Hamiltonian (1) has two distinct characters. On the average the centre of the wave packet follows a smooth classical trajectory but there is also a very fast oscillatory motion due to the interference of the positive and negative energy parts. In terms of density matrices, which we shall use below, we have the following. We may write, schematically, using the eigenkets $|e\rangle$ of Hamiltonian (1), an arbitrary density matrix as

$$\rho = \sum_{e,e'} |e\rangle\rho(e,e')\langle e'|. \tag{6}$$

Then the motion of this ρ will show *zitterbewegung* if and only if there are nonzero matrix elements $\rho(e, e')$ having the energies e, e' of opposite sign.

3. Subdynamics and transformation operators

The application to quantum mechanical systems of the ideas of subdynamics and superoperator transformations has been discussed in several papers by Prigogine and co-workers (Prigogine *et al* 1970a, b, Mandel 1970, Rae and Davidson 1971) so only a brief outline is given here. The starting point is the Von Neumann equation for the time evolution of a density matrix

$$\mathrm{i}\frac{\partial \rho}{\partial t} = [H, \rho] \equiv L\rho \tag{7}$$

where L is a linear operator acting on density matrices (hence the terminology superoperator) which are regarded as elements of some suitable linear space $\mathscr{L}$, for example the space of Hilbert–Schmidt operators. H is usually written as the sum of a 'solvable' part H_0 and a perturbation ϵH_1 and there is a corresponding decomposition of L into $L_0 + \epsilon L_1$: an important role is played by the nullspace projector for L_0, that is the projector P from $\mathscr{L}$ to the subspace $P\mathscr{L}$ of density matrices ρ for which $L_0\rho = 0$. In terms of the resolvent operator for L the solution of (7) is

$$\rho(t) = \frac{1}{2\pi \mathrm{i}} \int_C \frac{\mathrm{e}^{-\mathrm{i}zt}}{z - L} \rho(0)\, \mathrm{d}z = \mathrm{e}^{-\mathrm{i}tL}\rho(0)$$

with C the Bromwich contour from $+\infty$ to $-\infty$ parallel to the real z axis and above all singularities of the integrand. A careful separation of the contour into two parts provides, after a rather lengthy calculation, a projection operator Π which commutes with L and is such that the exact motion becomes the sum of two independent parts

$$\rho(t) = \mathrm{e}^{-\mathrm{i}tL}\Pi\rho(0) + \mathrm{e}^{-\mathrm{i}tL}(1 - \Pi)\rho(0). \tag{8}$$

The independent evolutions of the parts of ρ in the Π projected subspace (coherent part) and complementary subspace (incoherent part) constitute the subdynamics and this separation depends, among other things, on the projection operator P. It might be expected that, since P and Π are both projectors, a more direct relation can be found between them and this is indeed the case. It has been shown (Rae 1972a, b) that unitary superoperators V_α can be found such that V_α sends H into the nullspace of L_0: more precisely

$$V_\alpha(H_0 + \epsilon H_1) = H_0 + \epsilon K(\epsilon) \qquad L_0 K = 0 \tag{9}$$

and Π as obtained above is related to P by

$$\Pi = V_\alpha^\dagger P V_\alpha \tag{10}$$

the same Π resulting from any choice of V_α satisfying (9).

Now the FW transformation gives rise to a situation of exactly this type, for taking U given by (4) and defining the superoperator V by

$$V\rho = U\rho U^{-1} \tag{11}$$

we see immediately that V has the property (9). The FW transformation can therefore be used in (10) and (8) to find the independent subdynamics in this case. This is done in § 5.

The second aspect of Prigogine's theory which we wish to consider here is the transformation theory which has been developed for density matrices in the Π subspace. (Prigogine *et al* 1970a, b). Here the key role is played by a superoperator, generally denoted by χ, which acts from the subspace $P\mathscr{L}$ to $P\mathscr{L}$ and possesses an inverse in this subspace. Its definition following the original theory involves a rather lengthy chain of equations but is at root an expression in terms of P and L (see the form given in the

Appendix). For our purposes it is simpler to use the link that has been made (Rae 1972a, b, Rae and Davidson 1971) with the V superoperator of (9)

$$\chi = PV^{\dagger}P \tag{12}$$

where the V to be used in this expression is a particular case of (9) namely the unique superoperator V_{α} satisfying (9) and the subsidiary condition

$$P\frac{\partial V_{\alpha}}{\partial \epsilon}V_{\alpha}^{\dagger}P = 0. \tag{13}$$

It is easy to show that the V defined by (4) and (11) is exactly this unique choice (here the role of ϵ is played by $1/m$) and this will be done in the next section. Thus the FW transformation will enable us to calculate the superoperator χ. From among the transformed operators appearing in the theory we single out for examination later the transformed Hamiltonian H_{D} which is obtained from the original Hamiltonian by

$$H_{\mathrm{D}} = \chi^{-1}PH. \tag{14}$$

4. Superoperators associated with FW

In this section we produce explicit forms for superoperators associated with the Hamiltonian (1) and its transformation (4), considering in turn P, V, Π and χ.

The Hamiltonian (1) we put in the form

$$\frac{H}{m} = H_0 + \frac{1}{m}H_1 \qquad \text{with} \qquad H_0 = \beta, H_1 = \boldsymbol{\alpha}\,.\,\boldsymbol{p}$$

and $\epsilon = 1/m$ and write density matrices ρ in the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where a, b, c, d are 2×2 blocks. In this notation

$$L_0\rho \equiv [H_0, \rho] = 0 \Leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\Leftrightarrow b = c = 0.$$

Therefore the nullspace projector P is the block diagonal projector

$$P\rho = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}. \tag{15}$$

For future use we note the formula

$$P\rho = \tfrac{1}{2}\rho + \tfrac{1}{2}\beta\rho\beta. \tag{16}$$

The form of V given by (11) and (4) cannot usefully be simplified but following the remarks of § 3 we are obliged to check that V satisfies (13). It is enough to show $P(\partial V/\partial m)V^{\dagger}P = 0$ where m is taken as variable and $\boldsymbol{p}$ is fixed. From (4) one has directly

$$U(m+\delta) = U - \frac{m}{2E^2}\delta U + \frac{\delta}{2E}U^{\dagger} + \mathrm{O}(\delta^2)$$

where U, E etc are evaluated for mass m unless otherwise indicated. This gives

$$P\frac{\partial V}{\partial m}V^\dagger P\rho = \lim_{\delta\to 0}\frac{1}{\delta}P\{V(m+\delta)-V\}V^\dagger P\rho = \lim_{\delta\to 0}\frac{1}{\delta}P\{U(m+\delta)U^\dagger(P\rho)UU^\dagger(m+\delta)-P\rho\}$$

$$= P\left(-\frac{m}{2E^2}P\rho+\frac{1}{2E}(U^\dagger)^2P\rho+(P\rho)\frac{m}{2E^2}-(P\rho)U^2\frac{1}{2E}\right).$$

Since the P projector eliminates the odd parts it is easy to see from (4) that

$$P\{(U^\dagger)^2P\rho\} = \frac{m}{E}P\rho \qquad P\{(P\rho)U^2\} = (P\rho)\frac{m}{E}$$

so we obtain the required result

$$P\frac{\partial V}{\partial m}V^\dagger P = 0.$$

The projector Π is considered next. From (10) and (16) we have

$$\Pi\rho = U^\dagger(\tfrac{1}{2}U\rho U^\dagger+\tfrac{1}{2}\beta U\rho U^\dagger\beta)U = \tfrac{1}{2}\rho+\tfrac{1}{2}(U^\dagger\beta U)\rho(U^\dagger\beta U) = \frac{1}{2}\rho+\frac{1}{2}\frac{H}{E}\rho\frac{H}{E}. \tag{17}$$

Since $H^2 = E^2$ there follows at once from this:

$$\Pi^2 = \Pi \qquad [\Pi, L] = 0 \tag{18}$$

so Π is indeed a projector which commutes with L. Formula (17) is identical to the 'observable projection' of Pryce. The latter projection is closely related to time averaging (Pryce 1948, Crowther and Ter Haar 1971) which brings us very close to the original statistical mechanical motivation for the superoperator theory, so perhaps this connection is not surprising.

Finally, we consider the superoperator χ of equation (12) which we have shown above to be obtainable from the FW transformation.

We have

$$\chi\rho = PV^\dagger P\rho = \tfrac{1}{2}U^\dagger(P\rho)U+\tfrac{1}{2}\beta U^\dagger(P\rho)U\beta = \tfrac{1}{2}U^\dagger(P\rho)U+\tfrac{1}{2}U(P\rho)U^\dagger$$

so

$$\chi = \tfrac{1}{2}(V+V^\dagger)P. \tag{19}$$

This form shows quite clearly that χ^{-1} exists but there seems to be no simple explicit expression available for it.

5. Applications of the superoperators

We wish to consider two applications of the preceding section: firstly, a proof that the transformed Hamiltonian H_{D} of the superoperator theory is identical to the FW Hamiltonian and secondly, a demonstration that the two aspects of an electron's motion, the 'classical' part and the *zitterbewegung*, correspond exactly to the independent subdynamics in the two subspaces $\Pi\mathscr{L}$ and $(1-\Pi)\mathscr{L}$.

For the first part we have to show $\chi^{-1}PH = H_F$ with H_F given by (5). It is simplest to show instead that $\chi H_F = PH$. We have

$$\chi H_F = PV^\dagger P(\beta E) = \tfrac{1}{2}U^\dagger(E\beta)U + \tfrac{1}{2}\beta U^\dagger(E\beta)U\beta = \left(\frac{m+E}{2E} - \frac{p^2}{2E(E+m)}\right)E\beta$$
$$= m\beta = PH. \tag{20}$$

Although this proves what is required, the use of the superoperator V in this way solves the problem rather too simply in the sense that it effectively removes the element of surprise that two transformations of the Hamiltonian, H_D and H_F, should be identical although arising in totally different ways. It is instructive to perform the calculation without using V and relying on the older expressions for χ. Unfortunately such a calculation requires considerable familiarity with the superoperator theory and is in any case rather long so that a detailed account is neither possible nor desirable here. For those readers who are acquainted with the superoperator formalism an outline of the calculation is presented in an Appendix.

Turning now to the question of subdynamics we see that as Π commutes with L the time evolution in the subspace $\Pi\mathscr{L}$ is completely determined once we know which density matrices lie in $\Pi\mathscr{L}$. Accordingly we re-adopt the notation of (6) and use (17) to get

$$\Pi\rho = \frac{1}{2}\sum_{e,e'}|e\rangle\rho(e,e')\langle e'| + \frac{1}{2}\frac{H}{E}\sum_{e,e'}|e\rangle\rho(e,e')\langle e'|\frac{H}{E}$$
$$= \frac{1}{2}\sum_{e,e'}|e\rangle\rho(e,e')\langle e'| + \frac{1}{2}\frac{ee'}{|ee'|}\sum_{ee'}|e\rangle\rho(ee')\langle e'|.$$

Thus $\Pi\rho$ has matrix elements $\rho(e,e')$ if e and e' are both positive or both negative in sign, and zero otherwise. The space $\Pi\mathscr{L}$ therefore consists of those density matrices which do not couple positive and negative energy states and the space $(1-\Pi)\mathscr{L}$ consists of those which do. It follows from the remarks of § 2 that the motion in $\Pi\mathscr{L}$ is of a smooth classical type while the motion in $(1-\Pi)\mathscr{L}$ is the *zitterbewegung*. We illustrate this more explicitly by calculating the averages of the velocity operator $\mathrm{d}\boldsymbol{r}/\mathrm{d}t \equiv \boldsymbol{\alpha}$ with density matrices in each of these two subspaces.

In the following use is made of the obvious result

$$H\boldsymbol{\alpha} + \boldsymbol{\alpha}H = 2\boldsymbol{p}.$$

First we calculate the average with a density matrix $\rho \in \Pi\mathscr{L}$

$$\left\langle\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right\rangle = \mathrm{Tr}\left(\frac{1}{2}\boldsymbol{\alpha}\rho(t) + \frac{1}{2}\boldsymbol{\alpha}\frac{H}{E}\rho(t)\frac{H}{E}\right) = \mathrm{Tr}\left(\frac{1}{2}\boldsymbol{\alpha}\rho(t) + \frac{\boldsymbol{p}}{E}\rho(t)\frac{H}{E} - \frac{1}{2}\frac{H}{E}\boldsymbol{\alpha}\rho(t)\frac{H}{E}\right)$$
$$= \mathrm{Tr}\left(\frac{1}{2}\boldsymbol{\alpha}\rho(t) + \frac{H}{E}\frac{\boldsymbol{p}}{E}\rho(t) - \frac{1}{2}\boldsymbol{\alpha}\rho(t)\right) = \mathrm{Tr}\left(\frac{\boldsymbol{p}}{H}\rho(t)\right) = \mathrm{Tr}\left(\frac{\boldsymbol{p}}{H}\rho(0)\right)$$

which is just the average of the 'classical' velocity operator $\boldsymbol{p}/H$.

In a similar way if $\rho \in (1-\Pi)\mathscr{L}$ we have

$$\left\langle \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right\rangle = \mathrm{Tr}\left(\frac{1}{2}\boldsymbol{\alpha}\rho(t) - \frac{1}{2}\boldsymbol{\alpha}\frac{H}{E}\rho(t)\frac{H}{E}\right) = \mathrm{Tr}\left(\frac{1}{2}\boldsymbol{\alpha}\rho(t) - \frac{\boldsymbol{p}}{H}\rho(t) + \frac{1}{2}\boldsymbol{\alpha}\rho(t)\right) = \mathrm{Tr}\left\{\left(\boldsymbol{\alpha} - \frac{\boldsymbol{p}}{H}\right)\rho(t)\right\}$$

$$= \mathrm{Tr}\left\{\mathrm{e}^{\mathrm{i}Ht}\left(\boldsymbol{\alpha} - \frac{\boldsymbol{p}}{H}\right)\mathrm{e}^{-\mathrm{i}Ht}\rho(0)\right\} = \mathrm{Tr}\left\{\left(\boldsymbol{\alpha} - \frac{\boldsymbol{p}}{H}\right)\mathrm{e}^{-2\mathrm{i}Ht}\rho(0)\right\}.$$

This is exactly the oscillatory *zitterbewegung* term as calculated, for example, in Messiah (1966).

6. Extensions

The ideas of the preceding sections can be applied in other circumstances and here we mention two other examples.

In just the same way as occurs above for the free-particle Dirac equation, the FW transformation and *zitterbewegung* for the Klein–Gordon equation can be related to the superoperator theory. This is most easily seen if the Klein–Gordon equation

$$\left(\boldsymbol{p}^2 + m^2 - \frac{\partial^2}{\partial t^2}\right)\phi = 0$$

is rewritten, by taking $\partial\phi/\partial t$ as an independent variable, in its two component form (Feschbach and Villars 1958)

$$\mathrm{i}\frac{\partial\psi}{\partial t} = \eta\left(\frac{\boldsymbol{p}^2}{2m} + m\right)\psi + \nu\frac{\boldsymbol{p}^2}{2m}\psi \tag{21}$$

where ψ has two components and the matrices

$$\eta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \nu = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

play a role similar to β and $\boldsymbol{\alpha}$ of Dirac theory. In the space of functions ψ the scalar product is a little different from the usual one so that care has to be taken with the notions of adjointness and unitarity. However, one can find a unitary FW transformation

$$U = \frac{1}{2(mE)^{1/2}}\begin{pmatrix} E+m & E-m \\ E-m & E+m \end{pmatrix} \tag{22}$$

with $E = \sqrt{(m^2 + \boldsymbol{p}^2)}$ which takes the Hamiltonian in (21) to

$$H_{\mathrm{F}} = UHU^{\dagger} = \eta E. \tag{23}$$

Now we can proceed just as before. The results (16), (17), (18) and (19) all hold with η in place of β and the results of § 5 follow accordingly. The smooth motion lies in the subspace $\Pi\mathscr{L}$, the *zitterbewegung* in $(1-\Pi)\mathscr{L}$ and the transformed Hamiltonian is

$$H_{\mathrm{D}} = \chi^{-1}PH = \eta E \tag{24}$$

just as expected.

The foregoing can also be applied to cases where the Dirac or Klein–Gordon particle moves in an electromagnetic field. The main difference is that there is now no way, in general, to calculate a closed expression for the FW transformation and, as a result, the method employed above using the superoperator V now holds no computational advantage over the direct method used in the Appendix, this latter method being vastly more cumbersome in the free-particle case. If one considers the electromagnetic field as weak and calculates superoperators order by order, the results go through much as before. In particular, the calculation of $\chi^{-1}PH$ gives, term by term, the usual FW transformed Hamiltonian as calculated, for example, in Bjorken and Drell (1964). (For the lowest order terms see the Appendix, equation (A.8).)

Acknowledgments

The author wishes to thank Professors I Prigogine, L Rosenfeld, and R Balescu for discussions on the subject of this article.

Appendix. The direct calculation of H_{D} in the Dirac case

In the superoperator theory (Prigogine *et al* 1970a, b, Turner 1971) the transforming operator χ^{-1} has been calculated as

$$\chi^{-1} = P + \sum_{p,q:1}^{\infty} C_{pq} D^{(p)} C^{(q)} + \sum_{p,q,r,s} C_{pqrs} D^{(p)} C^{(q)} D^{(r)} C^{(s)} + \ldots \tag{A.1}$$

where

$$C_{pq} = q/(q+p) \qquad C_{pqrs} = -ps/(p+q)(p+q+r+s) \quad \text{etc} \tag{A.2}$$

and the labels p, q etc indicate the order in the perturbation parameter.

The operators D, C are defined as rather complicated expressions in terms of certain other operators $\psi, \mathscr{C}, \mathscr{D}$ and their derivatives, these being in turn obtainable from L and P. For an account of this see Prigogine *et al* (1969, 1970a, b) and George (1967): from now on we use various properties of these operators without further proof or comment. Since the operator ψ is nonzero even for the simple Hamiltonian (1) a general resummation of the series (A.1) is out of the question, but fortunately in calculating $\chi^{-1}PH$ important simplifications occur. The calculations are still lengthy so only an outline is given here.

First one can show quite simply

$$\begin{aligned} C^{(1)}H_0 &= \mathscr{C}^{(1)}H_0 = -\frac{1}{L_0}(1-P)L_1(m\beta) = \boldsymbol{\alpha}\,.\,\boldsymbol{p} \\ C^{(r)}H_0 &= 0 \qquad r > 1 \end{aligned} \tag{A.3}$$

and, more generally

$$C^{(1)}mf(\boldsymbol{p})\beta = \boldsymbol{\alpha}\,.\,\boldsymbol{p}f(\boldsymbol{p}) \qquad C^{(r)}mf(\boldsymbol{p})\beta = 0 \qquad r > 1 \tag{A.4}$$

for any scalar function f. Thus the first C on the right of terms in (A.1) behaves in a simple enough manner. Next we must consider the action of D, defined as

$$D = \sum_{m=0}^{\infty} \frac{1}{m!}(\overline{\Omega\psi})^m \frac{\partial^m \mathscr{D}}{\partial z^m}\bigg|_{z=0} \tag{A.5}$$

By some lengthy calculations one can show that derivatives of $\mathscr{D}$ acting on $\boldsymbol{\alpha}\,.\,\boldsymbol{p}$ always give either zero or terms of the form $g(\boldsymbol{p})\beta$, and that derivatives of ψ acting on this in turn give either zero or terms of the same form. But the ψ operator which implicitly stands at the left of $\overline{\Omega\psi}$ annihilates all such terms, the net effect being that only the term $m = 0$ of (A.5) contributes. Thus

$$D^{(1)}\boldsymbol{\alpha}\cdot\boldsymbol{p} = \mathscr{D}^{(1)}\boldsymbol{\alpha}\cdot\boldsymbol{p} = \frac{\boldsymbol{p}^2}{m}\beta$$
$$D^{(r)}\boldsymbol{\alpha}\cdot\boldsymbol{p} = 0 \qquad r > 1. \tag{A.6}$$

Now one may apply (A.4) and go through the steps again. The expression $\chi^{-1}PH$ reduces to the form

$$\chi^{-1}PH = (1 + C_{11}\mathscr{D}^{(1)}\mathscr{C}^{(1)} + C_{1111}\mathscr{D}^{(1)}\mathscr{C}^{(1)}\mathscr{D}^{(1)}\mathscr{C}^{(1)} + \ldots)PH.$$

Turner (1971) has shown that the coefficients C_{11}, C_{1111} etc are exactly the binomial coefficients $\binom{1/2}{n}$. Therefore

$$\chi^{-1}PH = \sum_{n:0}^{\infty}\binom{1/2}{n}[\mathscr{D}^{(1)}\mathscr{C}^{(1)}]^n PH = \sum_{n:0}^{\infty}\binom{1/2}{n}\left(\frac{\boldsymbol{p}^2}{m^2}\right)^n m\beta = \beta\sqrt{(m^2+\boldsymbol{p}^2)}$$

in accordance with (5).

In the case of an electromagnetic field the Dirac Hamiltonian is

$$H = m\beta + \boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}) + e\phi. \tag{A.7}$$

If we split this into $H_0 = m\beta$ and $V = \boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}) + e\phi$ we obtain the same L_0 and P as for free particles and find $PH = m\beta + e\phi$. The lowest order terms, in powers of $1/m$ and the field strength, coming from $\chi^{-1}PH$ are

$$\chi^{-1}PH = PH + \frac{1}{2}\mathscr{D}^{(1)}\mathscr{C}^{(1)}m\beta + \ldots = m\beta + e\phi + \frac{1}{2}PL\frac{1}{L_0}(1-P)\frac{1}{L_0}(1-P)Lm\beta + \ldots$$
$$= m\beta + e\phi + \frac{1}{2}PL\frac{1}{L_0}(1-P)\boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}) \tag{A.8}$$
$$= m\beta + e\phi + \frac{\beta}{2m}\{\boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A})\}^2 + \ldots.$$

References

Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
Crowther J M and Ter Haar D 1971 *Proc. K. Ned. Akad. Wet.* B **74** 241–50, 251–6
Feshbach H and Villars F 1958 *Rev. mod. Phys.* **30** 24–45
George Cl 1967 *Bull. Acad. R. Belg. Cl. Sci.* **53** 623–41
Mandel P 1970 *Physica* **50** 77–87
Messiah A 1966 *Quantum Mechanics* vol 2 (Amsterdam: North Holland)
Prigogine I *et al* 1969 *Physica* **45** 418–34
—— 1970a *Proc. Natn. Acad. Sci.* **65** 789–96
—— 1970b *Proc. Natn. Acad. Sci.* **66** 709–15
Pryce M H L 1948 *Proc. R. Soc.* A **195** 62–81
Rae J 1972a *Lett Nuovo Cim.* **3** 520–2
—— 1972b *Physica* to be published
Rae J and Davidson R 1971 *J. statist. Phys.* **3** 135–47
Turner J Wm 1971 *Physica* **51** 351–64