and p are in the hydrogen ground state, a rather formidable requirement. This being an essentially low-energy phenomenon, we would expect any deviations of the photon spin from unity to have a very small effect, even if the slope of the photon trajectory were as large as that of the vacuum Regge pole.

In conclusion, we would like to draw attention to a few points which naturally arise. For example, the electron should be a Regge pole; this could be tested, in principle, by high-energy Compton scattering or by large-angle pair production. The photon trajectory is of course in reality somewhat schizoid, consisting in some sense of an isotopic scalar and isotopic vector part, although of course $\alpha_{1V}(0)=\alpha_{1S}(0)=1$. It is possible that the ρ^0 and ω mesons are "consorts" of the photon. Such a possibility would make the phenomenological model of Gell-Mann and Zachariasen[18] (which couples these particles to the conventional electric current density) quite reasonable. Is there an even-signature "photon" family leading to bound states and resonances with even angular momentum? For example, could this be the so-called vacuum trajectory which would then be expected to be an isotopic schizoid also? If high-energy cross sections for particles and antiparticles should approach different constants, such a possibility would be suggestive.

It is hard to believe that we can ever acquire a deep understanding of the dynamical nature of isotopic spin unless we broaden the scope of the S-matrix approach to include electromagnetism.

[18] M. Gell-Mann and F. Zachariasen, Phys. Rev. **124**, 953 (1961).

PHYSICAL REVIEW VOLUME 128, NUMBER 5 DECEMBER 1, 1962

Extension of the Foldy-Wouthuysen Transformation

E. I. BLOUNT
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received July 23, 1962)

By use of Moyal's representation and a perturbation theory developed by the author, we remove the interband matrix elements of the Dirac Hamiltonian in the presence of external fields. The result is exact in terms of velocity, but appears as a series in the field strengths and their derivatives, which is evaluated to second order in the fields and first order in their first derivatives.

INTRODUCTION

IN 1950, Foldy and Wouthuysen[1] showed how one could derive the Pauli spin Hamiltonian from the Dirac equation by a unitary transformation. That is, they removed the matrix elements connecting positive and negative energy states. In the absence of external fields, this is no more than a transformation to eigenstates and was done exactly. When external fields are present, the problem is more complicated and could be done only by a series of successive approximations, amounting to an expansion in powers of $1/m$. The procedure was carried to second order and yielded a nonrelativistic result.

The author has recently developed a procedure for the removal of interband matrix elements in solid-state problems,[2] which can easily be specialized to deal with the Dirac Hamiltonian, if we consider the positive and negative energy states as each constituting a doubly degenerate band. The separation thus obtained has a much greater range of validity than that of FW, being valid for all energies. The result is obtained as a series in the field strengths and their derivatives, which we carry out to second order in the field strengths and first order in their first derivatives, thus obtaining all terms found by FW and a number of others.

Moyal has shown that the multiplication of two[3] operator functions of $\mathbf{p}$ and $\mathbf{x}$ may be written in the following form. If $\mathcal{P}=\mathfrak{N}\mathcal{O}$, then

$$\mathcal{P}(\mathbf{p},\mathbf{x})=\mathfrak{N}(\mathbf{p},\mathbf{x})\mathcal{O}(\mathbf{p},\mathbf{x})+\frac{i}{2}\left(\frac{\partial\mathfrak{N}}{\partial x_i}\frac{\partial\mathcal{O}}{\partial p_i}-\frac{\partial\mathfrak{N}}{\partial p_i}\frac{\partial\mathcal{O}}{\partial x_i}\right)$$
$$-\frac{1}{8}\left(\frac{\partial^2\mathfrak{N}}{\partial x_i\partial x_j}\frac{\partial^2\mathcal{O}}{\partial p_i\partial p_j}-\frac{2\partial^2\mathfrak{N}}{\partial p_i\partial x_j}\frac{\partial^2\mathcal{O}}{\partial p_j\partial x_i}\right.$$
$$\left.+\frac{\partial^2\mathfrak{N}}{\partial p_i\partial p_j}\frac{\partial^2\mathcal{O}}{\partial x_i\partial x_j}\right)+\cdots,\quad(\hbar=1)\quad(1)$$

where the multiplication on the right ignores the noncommutativity of $\mathbf{p}$ and $\mathbf{x}$. In this equation $\mathfrak{N}$, $\mathcal{O}$, $\mathcal{P}$ are all numerical functions of $\mathbf{p}$ and $\mathbf{x}$; to obtain the quantum-mechanical analogs it is only necessary to perform a Fourier transformation given in detail in reference 2 or to write out the power series for say, $\mathfrak{N}$, in powers of $\mathbf{p}$ and $\mathbf{x}$, being careful in each order to use all possible permutations of $\mathbf{p}$'s and $\mathbf{x}$'s which may then be treated as operators. In the application of (1) to the

[1] L. Foldy and S. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[2] E. I. Blount in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1962), Vol. 13.
[3] J. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

Dirac problem, it is only necessary to bear in mind that the operators are also 4×4 matrices. All products are matrix products and noncommutativity of matrix multiplication must be respected.

The advantage of this apparently rather cumbersome multiplication is that it can be viewed as an expansion in powers of $\partial/\partial\mathbf{x}$. Thus, if we are dealing with slowly varying functions of $\mathbf{x}$, the terms rapidly become smaller. In our case, since $\mathbf{x}$ appears only in the potentials, we get an expansion in the field strengths and their derivatives. A systematic perturbation theory based on this fact is given in Sec. III of reference 2. Here we shall proceed inductively.

Given the Hamiltonian

$$\mathcal{H}=c\boldsymbol{\alpha}\cdot(\mathbf{p}+\mathbf{a})+\beta mc^2+U, \tag{2}$$

where Dirac notation is followed, except that we have absorbed e/c into $\mathbf{a}$ and e into U and where $\mathbf{a}$ and U may be function of $\mathbf{x}$ and t, we wish to eliminate the interband elements, those involving $\boldsymbol{\alpha}$'s, ρ_1, ρ_2, etc., and obtain a Hamiltonian using only the unit matrix, the $\boldsymbol{\sigma}$'s and β. In the absence of $\mathbf{a}$ and U, we have the Hamiltonian

$$H=c\boldsymbol{\alpha}\cdot\mathbf{p}+\beta mc^2, \tag{3}$$

with eigenfunctions $e^{i\mathbf{p}\cdot\mathbf{x}}u_{n\mathbf{p}}$, where u_{np} is a spinor, independent of $\mathbf{x}$, and n labels, for the moment, the four eigenfunctions at $\mathbf{p}$. We define

$$S_{nn'}(\mathbf{p})=u_{n\mathbf{p}}{}^{\dagger}u_{n'0}, \tag{4}$$

which transforms the $u_{n0}{}^{\dagger}$ to $u_{n\mathbf{p}}{}^{\dagger}$. Then

$$S(\mathbf{p})HS^{\dagger}(\mathbf{p})=\beta E(\mathbf{p}), \tag{5}$$

$$E(\mathbf{p})=(m^2c^4+p^2c^2)^{1/2},$$

$$S(p)\mathbf{x}S^{\dagger}(p)=\mathbf{R}+\mathfrak{X}(\mathbf{p}), \tag{6}$$

where

$$\mathbf{R}\equiv i\partial/\partial\mathbf{p},$$

$$\mathfrak{X}_{nn'}(\mathbf{p})\equiv i\left(u_{n\mathbf{p}}{}^{*}\frac{\partial u_{n'\mathbf{p}}}{\partial\mathbf{p}}\right),$$

$$\mathfrak{X}(\mathbf{p})=-\frac{(\partial E/\partial\mathbf{p})\times\boldsymbol{\sigma}}{2(E+mc^2)}+\frac{ic\beta\boldsymbol{\alpha}}{2E}-\frac{ic(\boldsymbol{\alpha}\cdot\partial E/\partial\mathbf{p})\mathbf{p}}{2E_p(E_p+mc^2)}, \tag{7}$$

$$Sc\boldsymbol{\alpha}S^{\dagger}=\mathfrak{V}(\mathbf{p})$$

$$=\beta\frac{\partial E}{\partial\mathbf{p}}-c\boldsymbol{\alpha}+\frac{c(\boldsymbol{\alpha}\cdot\partial E/\partial\mathbf{p})\mathbf{p}}{(E_p+mc^2)}. \tag{8}$$

It is also convenient to define $\mathfrak{x}$ and $\mathbf{v}$ as the parts of $\mathfrak{X}$ and $\mathfrak{V}$ containing $\boldsymbol{\sigma}$ and β, and $\mathbf{X}$ and $\mathbf{V}$ as the parts involving $\boldsymbol{\alpha}$. The commutator of $\mathbf{x}$ and H yields

$$\mathbf{V}=-(1/\hbar)[\mathbf{X},\beta E], \tag{9}$$

as can be seen directly also. The explicit forms of $\mathfrak{X}$ and $\mathfrak{V}$ are taken from FW.

Now, after defining $\mathbf{a}$ in the future to be a function of $\mathbf{R}$, and $\mathfrak{p}\equiv\mathbf{p}+\mathbf{a}$, let us blindly write

$$\mathcal{H}'=S(\mathfrak{p})\otimes[H(\mathfrak{p})+U(\mathbf{R})]\otimes S^{\dagger}(\mathbf{p})-iS\otimes(\partial S^{\dagger}/\partial t), \tag{10}$$

where $\otimes$ indicates operator multiplication. $H(\mathfrak{p})+U(\mathbf{R})$ is equivalent to $\mathcal{H}$ before the transformation since $\mathbf{x}=\mathbf{R}$ in the Dirac representation. To form the operator product on the right we used (1). It is clear that we obtain terms involving

$$\frac{\partial a_i}{\partial R_j},\ \frac{\partial a_i}{\partial t},\ \frac{\partial U}{\partial R_i},\ \left(\frac{\partial a_i}{\partial R_i}\frac{\partial a_m}{\partial R_n}\right),\ \frac{\partial^2 a_i}{\partial R_j\partial R_l},\ \frac{\partial^2 a_i}{\partial t\partial R_j},\ \frac{\partial^2 I}{\partial R_i\partial R_j},$$

$$\left(\frac{\partial a_0}{\partial t}\frac{\partial a_j}{\partial R_l}\right),\ \text{and}\ \left(\frac{\partial U}{\partial R_i}\frac{\partial a_j}{\partial R_l}\right).$$

Similarly, we form the product $S\otimes S^{\dagger}$ and find that it is 1 only to zeroth order. Our program will then be to form $(1+g^{(1)})\otimes\mathcal{H}'\otimes(1+g^{(1)\dagger})$ where $g^{(1)}$ is a Hermitian operator chosen so that $(1+g^{(1)})\otimes S\otimes S^{\dagger}\otimes(1+g^{(1)\dagger})$ is unitary to first order; we then similarly chose $g^{(2)}$ to obtain an operator unitary to second order. We can then use perturbation theory to remove the interband terms of the resulting $\mathcal{H}$ through second order. We will then have a Hamiltonian with no interband matrix elements of less than third order. In principle this could be carried on indefinitely, but we will stop at this point. This procedure has been discussed in more detail in reference 2.

The terms involving only the first spatial derivatives of a_i, and their squares are entirely analogous to a problem discussed recently by the author[4] in the context of solids. These terms can be taken directly from that paper, and we will not repeat the steps in their evaluation. The other terms are evaluated in the Appendix. The net result is

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_1+\mathcal{H}_2+\mathcal{H}_{11}, \tag{11a}$$

$$\mathcal{H}_0=\beta E(\mathfrak{p})+U(\mathbf{R}), \tag{11b}$$

$$\mathcal{H}_1=-\tfrac{1}{2}\{(\mathbf{B}\times\mathbf{v})\cdot\mathfrak{X}\}-\tfrac{1}{4}\{\mathbf{B}(\mathbf{V}\times\mathfrak{X})\}-\mathfrak{E}\cdot\mathfrak{x}-\mathfrak{E}\cdot\mathbf{X}, \tag{11c}$$

$$\mathcal{H}_2=-\tfrac{1}{4}\{(\mathbf{B}\times\mathfrak{X})_l,(\partial\mathfrak{X}_i/\partial p_l)\mathfrak{E}_i\}$$

$$+\text{magnetic term}, \tag{11d}$$

$$\mathcal{H}_{11}=-\frac{1}{2}\frac{\partial\mathfrak{E}_i}{\partial R_j}\mathfrak{X}_i\mathfrak{X}_j-\frac{1}{24}\frac{\partial^2 a_i}{\partial R_j\partial R_l}\frac{\partial^3 E}{\partial p_i\partial p_j\partial p_l}$$

$$+\frac{1}{12}\frac{\partial B_m}{\partial R_l}\epsilon_{mij}\left(i\left[V_i,\frac{\partial X_i}{\partial p_l}\right]-(v_i\mathfrak{X}_j\mathfrak{X}_l+\mathfrak{X}_i\mathfrak{X}_jv_i)\right.$$

$$\left.+\{\mathfrak{X}_l\{\mathfrak{X}_i,\mathfrak{V}_j\}\}+\tfrac{1}{2}[[\mathfrak{X}_i,\mathfrak{X}_j],V_l]\right). \tag{11e}$$

Here, all functions $\mathbf{v}$, $\mathfrak{X}$, etc., are the objects as defined in Eqs. (7) and (8), not their transforms.

[4] E. I. Blount, Phys. Rev. **126**, 1636 (1962).

It remains only to remove the interband first-order terms; for the magnetic term this has already been done in reference 4, the electric term is trivial, and the cross term vanishes. The final result is

$$\begin{aligned}\mathfrak{H}=&\beta E(\mathfrak{p})+U(\mathbf{R})-\frac{c^2\mathbf{B}\cdot\boldsymbol{\sigma}\beta}{2E}\\&+\frac{\mathfrak{E}\cdot\mathbf{v}\times\boldsymbol{\sigma}\beta}{2(E+mc^2)}+\beta\frac{\mathfrak{E}^2c^2}{8E^3}-\beta\frac{(\mathfrak{E}\cdot\mathbf{v})^2}{8E^2}-\frac{\boldsymbol{\nabla}\cdot\mathfrak{E}c^2}{4E(E+mc^2)}\\&+\frac{(\mathbf{v}\cdot\boldsymbol{\nabla})\mathfrak{E}\cdot\mathbf{v}(2E^2+2Emc^2+m^2c^4)}{8E^2(E+mc^2)^2}-\frac{c^2(2E^2-m^2c^4)}{8E^3(E+mc^2)^2}\\&\times(\mathfrak{E}\times\mathbf{B})\cdot\mathbf{v}\beta+\frac{1}{12}(\mathbf{v}\cdot\boldsymbol{\nabla}\times\mathbf{B})\frac{2(E-mc^2)}{E^2(E+mc^2)}\\&-\frac{c^2}{2}\beta\left[\frac{(\mathbf{B}\times\mathbf{v})^2mc^2}{E^3(E+mc^2)}+\frac{(\mathbf{B}\cdot\mathbf{v})^2}{E(E+mc^2)^2}\right]\\&+\left[\beta\frac{c^2(\mathbf{B}\cdot\mathbf{v})^2}{4E(E+mc^2)^2}+\beta\frac{c^4(\mathbf{v}\times\mathbf{B})^2m^2c^4}{8E^3(E+mc^2)^2}\right]\\&+\frac{c^2}{16E^3}\beta[B^2c^2-(\mathbf{B}\times\mathbf{v})^2]\\&---\\&-\frac{1}{24}\frac{\partial^2a_i}{\partial R_j\partial R_l}\beta\frac{\partial^2E}{\partial p_i\partial p_j\partial p_l}\\&-\frac{c^2}{8E^3}\beta(B^2c^2-(\mathbf{B}\times\mathbf{v})^2).\end{aligned}\tag{12}$$

We now describe the meaning of the various terms:

(i) The last two terms are isolated from the others because they combine with the first term to make $\beta E(\mathfrak{p})$ into a more useful form $\beta E_g(\mathfrak{p})$. $E(\mathfrak{p})$ according to our previous statement becomes the quantum-mechanical operator obtained by weighing equally all permutations of the $\mathbf{p}$'s and $\mathbf{x}$'s in its series expansion as a function of $\mathbf{p}$ and $\mathbf{x}$. $E_g(p)$ becomes the quantum-mechanical operator obtained by substituting $\mathfrak{p}_i$ for p_i in the expansion of $E(\mathbf{p})$, again weighing all permutations of the $\mathfrak{p}_i$'s equally. $E_g(\mathfrak{p})$ is gauge invariant, whereas $E(\mathfrak{p})$ is not.

(ii) $U(\mathbf{R})$ speaks for itself.

(iii) $-c^2\mathbf{B}\cdot\boldsymbol{\sigma}\beta/2E$ is the magnetic moment term. Note that we obtain a g factor decreasing with energy which can also be written $-c^2\boldsymbol{\sigma}\cdot\mathbf{B}/m_t$, where m_t is the familiar transverse mass.

(iv) The fourth term is the spin-orbit coupling We recall that one would naively expect a term $-\mathfrak{E}\cdot\mathbf{v}\times\boldsymbol{\sigma}/2m$, by applying the Lorentz transformation to give an electric moment arising from the intrinsic magnetic moment. It is well known that this leads to an error by a factor of 2 commonly attributed to the "Thomas precession."[5] We observe that the factor of $\frac{1}{2}$ arises from the presence of (E_p+mc^2) rather than E_p. This result agrees with Thomas's expression [his Eq. (4.122), after dividing by β to make the left side $d\mathbf{w}/dt$] based on a classical electron with spin. As this is also true of the magnetic moment term, both first-order terms agree with Thomas's expression which requires only the assumption that the spin g factor is 2. For the higher order terms, Thomas's analysis would require further assumptions about the structure of the electron. It should also be observed that the high-energy form of the magnetic moment also can only be understood by taking account of the Thomas precession. It differs by a factor of $(mc^2/E_p)^2$ from the value that would be predicted by simple application of special relativity, which would in addition predict an anisotropic g factor.

(v) The terms in $\boldsymbol{\nabla}\cdot\mathfrak{E}$ and $(\mathbf{v}\cdot\boldsymbol{\nabla})\mathfrak{E}\cdot\mathbf{v}$ are due to the finite spread of the electron. When $\mathfrak{E}$ is entirely due to a potential we can regroup some terms as follows:

$$\begin{aligned}U(\mathbf{R})&-\frac{\mathfrak{E}\cdot\mathbf{v}\times\boldsymbol{\sigma}\beta}{2(E+mc^2)}-\frac{\boldsymbol{\nabla}\cdot\mathfrak{E}c^2}{4E(E+mc^2)}\\&+\frac{(\mathbf{v}\cdot\boldsymbol{\nabla})\mathfrak{E}\cdot\mathbf{v}(2E^2+2Emc^2+m^2c^4)}{8E^2(E+mc^2)^2}\\&=U\left(\mathbf{R}-\frac{\mathbf{v}\times\boldsymbol{\sigma}\beta}{2(E_p+mc^2)}\right)-\frac{c^2}{8E^2}\boldsymbol{\nabla}\cdot\mathfrak{E}+\frac{(\mathbf{v}\cdot\boldsymbol{\nabla})\mathfrak{E}\cdot v}{8E^2}.\end{aligned}\tag{13}$$

The first term represents the displacement of the center of the electron, while the last two describe the effect of its spread, being derived from the term $-\frac{1}{2}(\mathbf{X}\cdot\boldsymbol{\nabla})(\mathfrak{E}\cdot\mathbf{X})$. This spread is of the order of the Compton wavelength at low energies, and of the order of the de Broglie wavelength at high energies, where the Lorentz contraction also introduces a quadrupole moment, evidenced by the last term in (13). The reality and precise meaning of this spread can be seen by constructing a wave packet of positive (or negative) energy states, of momentum near $\mathbf{p}$. It is easily shown that such a packet cannot be more localized in space than c/E.

The terms discussed so far are the relativistic forms for the terms discussed by FW. The remaining terms, except for one, are of second order in the field strengths.

(vi) The terms quadratic in $\mathfrak{E}$ represent the polarizability of the electron, the fact that in an electric field its shape is distorted.

(vii) The terms quadratic in $\mathbf{B}$ are of several types. The first bracket consists of terms related to $-(\partial/\partial p_l)(\mathbf{B}\cdot\mathbf{m})\mathfrak{a}_l$, $(\mathfrak{a}\equiv\frac{1}{2}\mathbf{B}\times\mathfrak{x})$, $\mathbf{B}\cdot\mathbf{m}$ evaluated at $\mathfrak{p}+\mathfrak{a}$, rather than $\mathfrak{p}$, that is, they show that the electron

[5] L. H. Thomas, Phil. Mag. 3, 1 (1926); H. C. Corben and P. Stehle, *Classical Mechanics* (John Wiley & Sons, Inc., New York, 1950), p. 355.

is, in a sense, at $\mathbf{R}+\mathfrak{x}$. The second bracket consists essentially of terms due to the evaluation of $E(\mathfrak{p})$ at this point, as well as terms due to the second-order effect of the interband magnetic moment operator. The last bracket is similar to the standard atomic diamagnetism arising like the electric polarizability from the spread of the electron

(viii) The term in $\mathfrak{E}\times\mathbf{B}$ is attributable to the evaluation of $-\mathfrak{E}\cdot\mathfrak{X}$ at $\mathfrak{p}+\mathfrak{A}$, $(\mathfrak{A}=\frac{1}{2}\mathbf{B}\times\mathfrak{X})$.

(ix) The term in $\mathbf{v}\cdot\nabla\times\mathbf{B}$ is due to the evaluation of $\mathbf{B}$ at $\mathfrak{p}+\mathfrak{A}$, and to the noncommutativity of $\mathbf{B}$ and $\mathfrak{X}$ and $\mathfrak{V}$ as quantum-mechanical operators.

We now consider the representation of $\mathbf{x}$ as affected by the transformation we have made. To first order, we find easily

$$\mathbf{x}=\mathbf{R}+\mathfrak{X}(\mathfrak{p})+\tfrac{1}{2}\{(\partial\mathfrak{X}/\partial p_i)\mathfrak{A}_i\}+i[\mathfrak{X},\mathbf{Y}\cdot\mathfrak{E}]+\cdots, \quad (14)$$

where $\mathbf{Y}$ is the Hermitian operator with no intraband elements, which satisfies the equation

$$i[H,\mathbf{Y}\cdot\mathfrak{E}]=\mathfrak{E}\cdot\mathbf{X} \quad (15)$$

for all $\mathfrak{E}$. The dots indicate interband terms of first order, and higher order terms. The third term can be combined with the second to yield $\mathbf{X}(\mathfrak{p}+\mathfrak{A})$. The last term is due to the polarization of the electron by the electric field. Equation (14) takes the explicit form

$$\mathbf{x}=\mathbf{R}-\frac{\mathbf{v}\times\boldsymbol{\sigma}\beta}{2(E+mc^2)}+\frac{c^2(2E^2-m^2c^4)}{8E^2(E+mc^2)^2}\mathbf{B}\times\mathbf{v}\beta-\beta\times\left(\frac{\mathfrak{E}c^2}{E^3}-\frac{\mathfrak{E}\cdot\mathbf{v}\mathbf{v}}{E^3}\right)+\text{interband terms}+\text{higher order terms}. \quad (16)$$

We can now evaluate the intraband velocity from (12) and (16), obtaining the first order.

$$\mathbf{v}'=\mathbf{v}(\mathfrak{p})-\mathfrak{E}\times\boldsymbol{\Omega}+\frac{c^2\mathbf{B}\cdot\boldsymbol{\sigma}\beta\mathbf{v}}{2E^2}+\frac{\mathbf{v}c^2\mathbf{B}\times(\mathbf{v}\times\boldsymbol{\sigma})}{2E(E+mc^2)}+\cdots,$$
$$\boldsymbol{\Omega}\equiv\frac{2mc^2-E}{E^3}c^2\boldsymbol{\sigma}+\frac{c^2\boldsymbol{\sigma}\cdot\mathbf{v}\mathbf{v}}{2E(E+mc^2)}. \quad (17)$$

Thus, the velocity contains terms depending explicitly on the fields. The first of these terms is familiar in the nonrelativistic limit as the spin-orbit contribution to the velocity, but multiplied by a factor of 2. This results from the neglect of $\mathfrak{x}$ in the representation of $\mathbf{x}$ in the usual treatment of the Pauli Hamiltonian. The last two terms combine contributions of the form $(\mathbf{B}\times\mathbf{v})\times\boldsymbol{\Omega}$ with those of the form $+i[\mathbf{R}+\mathfrak{x},\mathbf{B}\cdot\mathbf{m}]$. The neglect of $\mathfrak{x}$ in the Pauli theory means that the velocity is not strictly correct, but in most cases we have to consider many electrons and a current given by an integral of the velocity over occupied states. The neglected terms can also be written $-i[\mathfrak{x},H]$; it is readily seen that for a steady-state system, the integral vanishes, so that for such a situation the usual treatment gives the total current correctly.

It would be tempting to try to find a simple semiclassical model of an electron which would behave like our transformed Dirac electron. So far as the Hamiltonian is concerned, we have seen that this is possible. When we try to understand the representation of $\mathbf{x}$, however, things are more complicated. In the first place, if we consider the term $\mathfrak{x}$, we see that while we could understand it on the basis of Thomas precession in the Hamiltonian, it also turns up in the absence of any external field in the expectation value of $\mathfrak{x}$. Similarly, the value of X^2 is smaller by a factor $(mc^2/E)^2$ than would be expected on the basis of a Lorentz transformation. These difficulties are all aspects of the fact that a Lorentz transformation cannot be represented by an intraband operator, but rather requires the use of $\boldsymbol{\alpha}$'s.[6] This statement shows clearly that any classical one-band picture is doomed to failure.

The last paragraph also points up the fact that it is not sufficient to specify a field-free Hamiltonian in order to determine a theory completely, but all other operators which can be varied, subject to internal consistency, must be given also. Thus, given the Dirac Hamiltonian

$$H=\beta mc^2+c\boldsymbol{\alpha}\cdot\mathbf{p},$$

we can choose $\mathbf{x}=\mathbf{R}+\mathbf{F}(\mathbf{p})$, where $\mathbf{F}$ is an arbitrary matrix function of $\mathbf{p}$. Dirac's choice $\mathbf{F}(\mathbf{p})=0$ represents a physical assumption. A different choice could, for example, have made $\mathbf{x}=\mathbf{R}$ in the FW representation; the velocity would have been $\mathfrak{V}=\mathbf{v}$, and the spin magnetic moment would have been zero. Dirac's choice is clearly preferred, but by experiment, not *a priori*.

Our interpretation of the FW representation does not differ significantly from theirs, except for a fairly trivial semantic detail. They ascribe $\mathfrak{X}$ to the "zitterbewegung" of the electron, whereas we would ascribe only $\mathbf{X}$ to this source, since $\mathfrak{x}$ actually represents a shift in the position. A wave packet made up of FW states will have its center shifted by $\langle\mathfrak{x}\rangle$ from the value given by the envelope, whereas X appears only in the spread of the packet.

The author's attention has been directed to the work of K. M. Case [Phys. Rev. **95**, 1323 (1954)] in which it is shown that in the case of a constant but not necessarily uniform magnetic field, the interband elements can be removed exactly to yield

$$\mathfrak{H}'=\beta(m^2c^4+c^2\mathfrak{p}^2-e\hbar c\boldsymbol{\sigma}\cdot\mathfrak{p})^{1/2}.$$

A straightforward computation shows that the terms in (12) independent of $\mathfrak{E}$ are the expansion of Case's form to the required order. It may be noted that Case was interested not only in Dirac particles but also in those of higher spin, for which he was unable to

[6] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1947), 3rd ed., p. 258.

generalize this result. Our procedure, while not exact, can readily be generalized to arbitrary spin.

APPENDIX

The evaluation of (10) divides naturally into several parts. The terms involving only the first spatial derivatives of $\mathbf{a}_i$ are to be found in reference 4 and we merely quote the result:

$$\begin{aligned}\mathfrak{H}_{cm}=&E(\mathfrak{p})-\tfrac{1}{2}\{(\mathbf{B}\times\mathbf{v})\cdot\mathfrak{X}\}-\tfrac{1}{4}\{\mathbf{B}\cdot(\mathbf{V}\times\mathfrak{X})\}\\&+\tfrac{3}{8}\{\mathfrak{A}\{\mathfrak{A}\{\mathfrak{A}_j,\alpha_{ij}\}\}+\frac{\mathfrak{A}^2}{2m}+\frac{1}{2}\left\{v_j\left\{\mathfrak{A}_i,\frac{\partial\mathfrak{A}_j}{\partial p_i}\right\}\right\}\\&+\frac{1}{4}\left[[\mathfrak{A}_j,v_i],\frac{\partial\mathfrak{A}_i}{\partial p_j}\right]+\frac{1}{8}\left[\frac{\partial\mathfrak{A}_j}{\partial p_i},\left[\frac{\partial\mathfrak{A}_i}{\partial p_j},E\right]\right]\\&-\frac{1}{8}\left\{\mathfrak{A}_i,\frac{\partial}{\partial p_i}\mathbf{B}\cdot\{\mathbf{V}\times\mathbf{X}\}\right\},\end{aligned}\tag{A1}$$

where $\mathfrak{A}=\frac{1}{2}\mathbf{B}\times\mathfrak{X}$.

The next group consists of terms from $S\otimes H\otimes S^\dagger$ involving second derivatives of $\mathbf{a}_i$:

$$\begin{aligned}\mathfrak{H}_{dm}=&-\frac{1}{8}\frac{\partial^2 a_i}{\partial R_j\partial R_l}[S_iH_{jl}S^\dagger+S_iH_jS_l^\dagger+SH_iS_{jl}^\dagger\\&-S_lH_iS_j^\dagger+S_iH_lS_j^\dagger+S_iHS_{jl}^\dagger+S_{jl}H_iS^\dagger-S_jH_iS_l^\dagger\\&+S_{jl}HS_i^\dagger+S_jH_lS_i^\dagger+S_lH_jS_i^\dagger+SH_{jl}S_i^\dagger\\&+S_jS_i^\dagger(SHS^\dagger)_l+(SHS^\dagger)_lS_iS_j^\dagger\\&-S_iS_j^\dagger(SHS^\dagger)_l-(SHS^\dagger)_lS_jS_i^\dagger].\end{aligned}\tag{A2}$$

The subscripts indicate differentiation. The last four terms arise not from $S\times H\times S^\dagger$ but from $g^{(1)}\otimes SHS^\dagger+SHS^\dagger\otimes g^{(1)}$, where $(1+g^{(1)})$ is the additional transformation used in reference 4 to make S unitary through first order. Such terms as well as $g^{(2)}$ are included in (A1). If we replace H by 1 in (A2) we do not obtain zero, showing that additional terms in $\partial^2 a_i/\partial R_j\partial R_l$ are also required to make S unitary through second order. First, however, we split (A2) into two parts, the first symmetrical in the interchange of indices, the second the remainder. On adding the terms required for unitarity in the form of a $g^{(2)}$, we obtain

$$\begin{aligned}\mathfrak{H}_{dm}=&-\frac{1}{24}\frac{\partial^2 a_i}{\partial R_j\partial R_l}\frac{\partial^2 E}{\partial p_i\partial p_j\partial p_l}+\frac{1}{8}\frac{\partial^2 a_i}{\partial R_j\partial R_l}[[\mathfrak{X}_i,\mathfrak{X}_j],v_l]\\&-\frac{1}{12}\frac{\partial^2 a_i}{\partial R_j\partial R_l}\left(-i\left[v_i,\frac{\partial\mathfrak{X}_j}{\partial p_l}\right]+i\left[v_j,\frac{\partial\mathfrak{X}_i}{\partial p_l}\right]-v_i\mathfrak{X}_j\mathfrak{X}_l\right.\\&+v_j\mathfrak{X}_i\mathfrak{X}_l-\mathfrak{X}_l\mathfrak{X}_jv_i+\mathfrak{X}_l\mathfrak{X}_iv_j+2\mathfrak{X}_i\mathfrak{V}_j\mathfrak{X}_l\\&\left.+2\mathfrak{X}_l\mathfrak{V}_j\mathfrak{X}_i-4\mathfrak{X}_j\mathfrak{V}_i\mathfrak{X}_l\right),\end{aligned}\tag{A3}$$

where we have used (6) and (8). Rearrangement yields

$$\begin{aligned}&-\frac{1}{24}\frac{\partial^2 a_i}{\partial R_j\partial R_l}\frac{\partial^2 E}{\partial p_i\partial p_j\partial p_l}-\frac{1}{12}\frac{\partial B_m}{\partial R_l}\epsilon_{mij}\left(i\left[\mathfrak{V}_i,\frac{\partial\mathfrak{X}_j}{\partial p_l}\right]\right.\\&\left.-\{\{\mathfrak{X}_i,\mathfrak{V}_j\}\mathfrak{X}_l\}+v_i\{\mathfrak{X}_j\mathfrak{X}_l\}\right)+\text{interband terms.}\end{aligned}\tag{A4}$$

The next terms involve U:

$$\begin{aligned}\mathfrak{H}_u=&U(\mathbf{R})+(i/2)(SU_iS_i^\dagger-S_iU_iS^\dagger)\\&-\tfrac{1}{8}(SU_{ij}S_{ij}^\dagger+S_{ij}U_{ij}S^\dagger-2S_iU_{ij}S_j^\dagger)\\&-\tfrac{1}{4}[\partial a_j/\partial R_l-\partial a_l/\partial R_j](S_jU_iS_{il}^\dagger-S_{ij}U_iS_l^\dagger)\\&+\tfrac{1}{8}[\partial a_j/\partial R_l-\partial a_l/\partial R_j]\\&\times\{S_jS_l^\dagger,(SU_iS_i^\dagger-S_iU_iS^\dagger)\},\end{aligned}\tag{A5}$$

the last term being due to $g^{(1)}$. Equation (A5) simplifies to

$$\begin{aligned}\mathfrak{H}_u=&U+U_i\mathfrak{X}_i-\tfrac{1}{2}U_{ij}\mathfrak{X}_i\mathfrak{X}_j\\&+\tfrac{1}{4}\{(\mathbf{B}\times\mathfrak{X})_l,(\partial\mathfrak{X}_i/\partial p_l)U_i\}.\end{aligned}\tag{A6}$$

The terms involving $\partial S/\partial t$ are handled similarly and their only effect is to change U_i in (A6) to the electric face $\mathfrak{E}_i=-U_i-(l/c)(\partial a_i/\partial t)$.

All the final expressions obtained here are also valid for electrons in crystals provided $\mathfrak{V}$, $\mathfrak{X}$, etc., are given their appropriate meanings as in reference 4.