In: *Journal of Mathematical Physics*, **16** 573–584 (1975).

Consistency in the formulation of the Dirac, Pauli, and Schrödinger theories

R. Gurtler and D. Hestenes

Abstract. Properties of observables in the Pauli and Schrödinger theories and first order relativistic approximations to them are *derived* from the Dirac theory. They are found to be *inconsistent* with customary interpretations in many respects. For example, failure to identify the "Darwin term" as the s-state spin-orbit energy in conventional treatments of the hydrogen atom is traced to a failure to distinguish between charge and momentum flow in the theory. Consistency with the Dirac theory is shown to imply that the Schrödinger equation describes not a spinless particle as universally assumed, but a particle in a spin eigenstate. The bearing of spin on the interpretation of the Schrödinger theory discussed. Conservation laws of the Dirac theory are formulated in terms of relative variables, and used to derive virial theorems and the corresponding conservation laws in the Pauli-Schrödinger theory.

Introduction

In quantum mechanics three different equations are widely used to describe the motion of a single electron, the Schrödinger, Pauli, and Dirac equations. Each of these equations must be supplemented by physical assumptions which prescribe how to calculate observables from the electron wavefunction. The three wave equations are intimately related; the Pauli equation being an approximation to Dirac equation for small electron velocities, while the Schrödinger equation approximates the Pauli equation by neglecting magnetic interactions of the spin.

Obviously, the observables associated with the three equations should be related to one another by the same approximations. In fact, however, quite a few inconsistencies in this regard are to be found in the literature. Consider, for example the usual expressions for probability density ρ and probability current in the Schrödinger theory,

$$\rho = \Psi^\dagger\Psi\,, \tag{1.1a}$$

$$\rho u_k = -\frac{i\hbar}{2m}\{\Psi^\dagger\partial_k\Psi - \partial_k\Psi^\dagger\Psi\} - \frac{e}{mc}\,A_k\Psi^\dagger\Psi\,. \tag{1.1b}$$

In the Schrödinger theory, Eq. (1.1b) plays a triple role; besides the probability current ρu_k it determines the charge current $e\rho u_k$ associated with a charge density $e\rho$ and a kinetic momentum density $m\rho u_k$ associated with a mass density $m\rho$.

In the Pauli theory the same expressions (1.1a,b) are usually used for probability density and current, Ψ being understood as the two component Pauli wavefunction instead of the Schrödinger wavefunction (e.g., Ref. 1). In both the Schrödinger and Pauli theories the wave-equation implies the conservation law

$$\partial_t\rho + \partial_k(\rho u_k) = 0\,. \tag{1.1c}$$

Though $e\rho$ is still interpreted as charge density, the Pauli theory differs from Schrödinger theory in that $e\rho u_k$ must be supplemented by a "spin magnetization current" $c\nabla\times\mathbf{m}$ to get the total charge current

$$\mathbf{j} = e\rho\mathbf{u} + c\nabla\times\mathbf{m}\,, \tag{1.2a}$$

where

$$m_k = \frac{e\hbar}{2mc}\,\Psi^\dagger\sigma_k\Psi \tag{1.2b}$$

and the σ_k are the usual Pauli matrices. The fact that (1.2a) is the correct expression for the charge current has been established for a long time,[2] though its importance seems to be frequently unappreciated.

Following history, textbooks show how to get from the Schrödinger theory to the Pauli theory by heuristic arguments (e.g., Ref. 1). Reversing the procedure, the Schrödinger theory can be derived rigorously (rather than heuristically) from the Pauli theory, with some consequences that seem to have been completely overlooked. When the magnetic field is small or zero, the Pauli wave equation is identical to the Schrödinger equation and possesses solutions of the form

$$\Psi = \begin{pmatrix} \phi \\ 0 \end{pmatrix} . \tag{1.3}$$

So from (1.1a) we have $\rho = \Psi^\dagger\Psi = \phi^\dagger\phi$ and from (1.1b) we get an expression for ρu_k with ϕ replacing Ψ. Since ϕ is a complex function satisfying the Schr̈odinger equation, and since its relation to the probability density and current has been derived, it would seem that we have arrived at the Schrödinger theory.

But here's the rub! we must take account of fact that Eq. (1.3) means that the electron is in an eigenstate of the spin. So what we have proved is that the *Schrödinger theory is identical to the Pauli theory when the electron is in an eigenstate of the spin.* Of course, this is at variance with the usual view that the Schrödinger theory describes a particle without spin, but it is a rigorous consequence of requiring that the theory be derivable from the Pauli theory. The difference is important! Though Eq. (1.3) yields some of the usual features of the Schrödinger theory, it also implies nonvanishing values for (1.2b), specifically, if σ_3 is diagonal as usual, then

$$m_3 = \frac{e\hbar}{2mc}\,\phi^\dagger\phi = \frac{e\hbar}{2mc}\,\rho\,, \qquad m_1 = m_2 = 0\,. \tag{1.4}$$

Equation (1.4) leads to nonvanishing values for the magnetization current $\nabla\times\mathbf{m}$. Hence, the expression for the charge current $\mathbf{j}$ given by (1.2a) does not reduce to $e\rho\mathbf{u}$ when the passage to the Schrödinger theory is made. That is, the usual expression for charge current in the

Schrödinger theory obtained by multiplying (1.1b) by the charge e is *inconsistent* with the assumption that the Schrödinger theory is derivable from the Pauli theory To put it bluntly, everyone to date has been using the wrong expression for charge current density in the Schrödinger theory. Of course there is no way that this error could be revealed directly by experiment, because the only direct experimental means of testing for the existence of a magnetization current is by introducing a magnetic field. But in that case everyone knows enough to discard the Schrödinger theory and use the Pauli or Dirac theories however, the existence of a magnetization current has important bearing on the interpretation of the Schr̈odinger theory even in the absence of a magnetic field For instance, it implies that there is a nonvanishing charge current in the s-states of hydrogen, eliminating one of the reputedly fundamental differences between the Schrödinger and Bohr theories. It also leads to the conclusion that the appearance of complex number in the Schrödinger theory is inseparably related to the existence of the spin, the factor $i\hbar$ being significant in the theory only because $\frac{1}{2}i\hbar$ is an eigenvalue of the matrix representing the spin. This is difficult to reconcile with conventional interpretations of the uncertainty principle.

Though it is supported experimentally, the expression (1.2a) for the charge current was originally introduced into the Pauli theory as an *ad hoc* assumption. So it is important to know that it can be justified on deeper theoretical grounds. The charge current in the Dirac theory is given by the well-known expression

$$j^\mu = e\,\overline{\Psi}\gamma^\mu\Psi\,, \tag{1.5}$$

where Ψ now the four component Dirac wavefunction. The Dirac equation implies the conservation law

$$\partial_\mu j^\mu = 0\,, \tag{1.6}$$

as well as the decomposition

$$j^\mu = \frac{e}{mc}\,k^\mu + \partial_\nu M^{\mu\nu}\,, \tag{1.7a}$$

where

$$k^\mu = \frac{i\hbar}{2}\{\,\overline{\Psi}\partial^\mu\Psi - (\partial^\mu\,\overline{\Psi})\Psi\} - \frac{e}{mc}\,A^\mu\,\overline{\Psi}\Psi \tag{1.7b}$$

is the so-called Gordon current,[3] and

$$M^{\mu\nu} = j^{\mu} = \frac{ie\hbar}{2mc}\,\Psi\tfrac{1}{2}[\gamma^{\nu},\gamma^{\mu}]\Psi \tag{1.7c}$$

is interpreted as the magnetic moment density due to the electron spin. In the nonrelativistic limit, the time and space components of (1.7b) reduce, respectively, to (1.1a) and (1.1b), while the nonvanishing components of (1.7c) are given in (1.2b), so the space components of (1.7a) reduce exactly to Eq. (1.2a). Thus, the expression (1.2a) for the charge current in the Pauli theory is fully justified by the requirement of consistency with the Dirac theory. It is important to realize that the particular combination of currents in (1.7a) is a consequence of the Dirac equation, whereas the limiting result (1.2a) is not a consequence of the Pauli equation.

The interpretation of the Dirac current (1.5) as a charge current is well established experimentally. But the Dirac current is also interpreted as a probability current by dropping the charge e. Here we run into another inconsistency with the Schrödinger theory, for the Schrödinger current (1.1b), which is supposed to be a probability current, corresponds to the Gordon current (1.7b) and not, as we have seen, to the Dirac current One way to resolve this difficulty might be to identify the Gordon current as the probability current in the Dirac theory. The required conservation law

$$\partial_{\mu} k^{\mu} = 0 \tag{1.8}$$

is indeed satisfied, but then new problems of normalization and interpretation arise in the Dirac theory. The only alternative is to conclude that (1.2a) rather than (1.1b) determines the correct probability current as well as the charge current. This does not mean that the Schrödinger current should be dispensed with. Indeed, comparison with the Dirac theory, it can be shown to be proportional to the momentum density (1.2a) suggests it can be interpreted l . convective charge

The main objective of this paper is to establish a consistent identification of observables in Dirac, Pauli, and Schrödinger theories. This will be accomplished by beginning with the formulation of the Dirac theory in terms of local observables as given in Refs. 4 and 5, and obtaining the corresponding formulation of the Pauli and Schrödinger theories as limiting cases. The unusual formulation of quantum theory employed here is fully equivalent mathematically to the conventional one. However, it brings to light certain problems in physical interpretation which, as already argued in Ref. 4, may require for their resolution some modification of current theory. No such modification of quantum theory will be attempted here. But we cannot resist expressing the opinion that the Dirac theory is best interpreted as describing statistical ensemble of particle motions and pointing out from time to time how this may help the understanding of mathematical relations in the theory. Though some of the unusual physical interpretations we suggest are open to dispute and hopefully at some time to experimental test, the mathematical steps alone show what is required to establish consistency among the Dirac, Pauli, and Schrödinger theories.

A formulation of the Dirac theory in terms of local observables like the one given in Ref. 4 is sometimes called a "hydrodynamic formulation" of quantum theory. Hydrodynamics provides a ready-made terminology for the description of continuous distributions and flows of mechanical quantities such as energy; momentum, angular momentum, and charge; as such it is useful in quantum theory, but it should be understood that the use of hydrodynamic terminology does not imply that any classical model or interpretation has been presumed. A hydrodynamic formulation of Schrödinger theory was first given by Madelung[7]: it has been discussed since by numerous authors, recently, for example, by Wilhelm.[8] Complete hydrodynamic formulations of the Pauli and Dirac theories were first given by Takabayasi,[9] though other authors, notably Costa de Beauregard,[10] achieved partial results earlier. These formulations of the Schrödinger, Pauli, and Dirac theories, though fully consistent with more conventional formulations of quantum theory are inconsistent with one another in their identifications of observables. In Ref. 6 it was shown that if the Schrödinger theory is regarded as an approximation to the Pauli theory, then it necessarily contains spin (albeit in a degenerate form). Her we show how the identification of observables in Ref. 5 must be adjusted to be consistent with the more fundamental formulation of the Dirac theory in Refs. 4 and 5.

Section 2 obtains the Pauli theory as the nonrelativistic approximation to the Dirac theory and discusses relativistic corrections. The usual physical interpretation of these results is held to be incorrect because of insufficient attention to the identification of observables, especially failure to bring the nonrelativistic limit of the Dirac (charge) current into the discussion and distinguish it from the momentum density. The Darwin term in the energy is proved to be a spin-orbit energy for s-states in exact accordance with the original argument of Thomas.

Section 3 summarizes the definitions and interrelations of observables in the Pauli-Schrödinger theory which are *required* for the sake of consistency with the Dirac theory. Some implications of the consistency requirement for the Schrödinger hydrogen atom are pointed out to show that the conventional interpretation of the Schrödinger theory must be drastically revised, but no attempt is made to carry any such revision to completion.

Section 4 expresses the hydrodynamic equations of the Dirac theory in terms of relative observables, uses them to derive a virial theorem, and obtains their non-relativistic limit.

The nomenclature and results of Refs. 4, 5, and 11 are used throughout this paper. The reader is advised to become familiar especially with Ref. 5 before attempting to follow the arguments here in any detail.

2. Nonrelativistic Approximations to the Dirac Theory

In the literature two methods have been widely used to generate nonrelativistic approximations to the Dirac theory, namely, separation of the Dirac wavefunction into large and small components,[12,13,14] and the F-W transformation.[15] To facilitate comparison with our approach, we translate the first of these methods into multivector language. Then we criticize the physical interpretation usually accorded to the method and give reasons for interpreting it differently. Our arguments also have bearing on the F-W transformation and suggest rather different mathematical methods for generating relativistic corrections, but we do not pursue either of these points in any detail. Our objective is only to show in multivector language how the Pauli equation and relativistic corrections to it can be obtained from the Dirac equation and provided with a consistent physical interpretation.

The definition of electron energy in the Dirac theory differs from the definition in the nonrelativistic theories by including the rest energy. We can remove the rest energy while retaining the definition of the energy in terms of the wavefunction by changing the wave equation with the transformation

$$\psi \to \psi \exp\{-i\boldsymbol{\sigma}_3 mc^2 t/\hbar\}\,, \tag{2.1}$$

whereupon the Dirac equation (3.11) of Ref. 5 becomes

$$\hbar\Box\psi i\boldsymbol{\sigma}_3 + mc\psi\gamma_0 = mc\gamma_0\psi + \frac{e}{c}\,A\psi\,. \tag{2.2}$$

To express (2.2) in terms of relative variables we multiply it by $c\gamma_0$ and, recalling the definitions (6.4) and (6.22) of Ref. 5, obtain

$$\hbar(\partial_t + c\nabla)\psi i\boldsymbol{\sigma}_3 = mc^2(\psi^* - \psi) + (V - e\,\mathbf{A})\psi\,, \tag{2.3}$$

where

$$\psi^* = \gamma_0\psi\gamma_0\,. \tag{2.4}$$

The wavefunction ψ can be expressed as the sum of an even part ψ_e and an odd part ψ_o, that is,

$$\psi = \psi_e + \psi_o\,. \tag{2.5a}$$

where distinction between "even" and "odd" is best made by the equation

$$\psi^* = \psi_e - \psi_o\,. \tag{2.5b}$$

As is easily shown by the method of Appendix A in Ref. 4, this separation of ψ into even and odd multivector parts is exactly equivalent to the usual separation of the wavefunction into large and small components in the matrix version of the Dirac theory, but we shall see only later what this separation means physically. After substituting (2.5) into (2.3) and separately equating even and odd parts, we obtain the coupled equations:

$$\hbar\partial_t\psi_e i\boldsymbol{\sigma}_3 = \big\{\hbar\nabla\psi_o i\boldsymbol{\sigma}_3 + \frac{e}{c}\,\mathbf{A}\psi_o\big\} + V\psi_e\,, \tag{2.6a}$$

$$\hbar\partial_t\psi_o i\boldsymbol{\sigma}_3 - V\psi_o + 2mc^2\psi_o = -c\big\{\hbar\nabla\psi_o i\boldsymbol{\sigma}_3 + \frac{e}{c}\,\mathbf{A}\psi_e\big\}\,. \tag{2.6b}$$

Equations (2.6) can be solved to lowest order by neglecting the first two terms in the left of (2.6b) in relation to the third, yielding

$$\psi_o = -\frac{1}{2mc}\Big\{\hbar\nabla\psi_e i\boldsymbol{\sigma}_3 + \frac{e}{c}\,\mathbf{A}\psi_o\Big\}\,. \tag{2.7a}$$

If we use this to eliminate ψ_o from (2.6a), then after expanding, simplifying and using the identity

$$\mathbf{A}\psi_e + \nabla(\mathbf{A}\psi_e) = (\nabla\mathbf{A}\cdot\nabla)\psi_e = (i\mathbf{B} + 2\mathbf{A}\cdot\nabla)\psi_e\,,$$

we arrive at the Pauli equation

$$\begin{aligned}\hbar\,\partial_t\psi_e i\boldsymbol{\sigma}_3 = \frac{1}{2mc}\Big\{-\hbar^2\nabla^2 + \frac{e^2}{c^2}\,\mathbf{A}^2\Big\}\psi_e \\ + \frac{e\hbar}{2mc}\Big\{i\mathbf{B} + 2\mathbf{A}\cdot\nabla\Big\}\psi_e i\boldsymbol{\sigma}_3 + V\psi_e\,.\end{aligned} \tag{2.7b}$$

For readers who are still not completely at home with the multivector algebra used here, we prove that (2.7) is equivalent to the usual matrix form of the Pauli equation. This is most easily done by replacing each vector $\boldsymbol{\sigma}_k$ in (2.7) by a corresponding Pauli matrix σ_k according to the rules

$$\begin{aligned}\boldsymbol{\sigma}_1 \to \sigma_1 &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\,,\\ \boldsymbol{\sigma}_2 \to \sigma_2 &= \begin{pmatrix}0 & -i'\\ i' & 0\end{pmatrix}\,,\\ \boldsymbol{\sigma}_3 \to \sigma_3 &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\,.\end{aligned} \tag{2.8}$$

The i' here is a mathematical square root of -1 with no geometrical interpretation; however, i' multiplied by the unit matrix is a matrix representation of the pseudoscalar i, a fundamental geometrical entity. This follows from (2.8), thusly:

$$i = \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 \to \sigma_1\sigma_2\sigma_3 = i'\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\,.$$

By (2.8) ψ_e corresponds to a matrix ψ'_e which becomes a column matrix Ψ_e by operating on an eigenmatrix u_1 of σ_3; so we make the correspondence

$$\psi_e \to \psi'_e\,, \quad \Psi_e \to \psi'_e u_1 \qquad \text{where} \qquad u_1 = \begin{pmatrix}1\\ 0\end{pmatrix}\,. \tag{2.9a}$$

From (2.8) we also have

$$\mathbf{B} = \boldsymbol{\sigma}_i B^i \to \sigma_i B^i \equiv \sigma\cdot\mathbf{B}\,. \tag{2.9b}$$

Hence regarding (2.7) as a matrix equation and multiplying it on the right by u_1, we get, by (2.8) and (2.9), the Pauli equation in its usual matrix form

$$\begin{aligned}i'\hbar\,\partial_t\Psi_e = \frac{1}{2m}\Big\{-\hbar^2\nabla^2 + \frac{e^2}{c^2}\,\mathbf{A}^2\Big\}\Psi_e \\ + \frac{e\hbar}{2mc}\Big\{-\sigma\cdot\mathbf{B} + 2i'\mathbf{A}\cdot\nabla\Big\}\Psi_e + V\Psi_e\,.\end{aligned} \tag{2.10}$$

Now we return to the multivector formalism and the coupled equations (2.6a,b).

To obtain higher order approximations to (2.6a,b) in a systematic way and to discuss the usual physical interpretation of the results, it is convenient to introduce operators $\hat{K}$, $\hat{\mathbf{p}}$, and $\hat{p}_k$ defined by

$$\hat{K}\psi \equiv \hbar\,\partial_t\psi i\boldsymbol{\sigma}_3 - V\psi\,, \tag{2.11}$$

$$\hat{\mathbf{p}}\psi \equiv -\hbar\nabla\psi i\boldsymbol{\sigma}_3 - \frac{e}{c}\,\mathbf{A}\psi\,, \tag{2.12a}$$

$$\hat{p}_k\psi \equiv -\hbar\,\partial_k\psi i\boldsymbol{\sigma}_3 - \frac{e}{c}\,\mathbf{A}_k\psi \equiv \boldsymbol{\sigma}_k\cdot\hat{p}\psi\,. \tag{2.12b}$$

The Pauli equation (2.7b) can then be written

$$\hat{K}\psi_e = \frac{1}{2m}\hat{\mathbf{p}}^2\psi_e = \frac{1}{2m}\hat{p}^2\psi_e - \frac{e\hbar}{2mc}\,\mathbf{B}\psi_e\boldsymbol{\sigma}_3\,, \tag{2.13}$$

where we have used the obvious operator notation

$$\hat{p}^2 \equiv \hat{\mathbf{p}}\cdot\hat{\mathbf{p}} = \hat{p}_k\cdot\hat{p}_k\,. \tag{2.14}$$

Equations (2.6a,b) can be put in the form

$$\hat{K}\psi_e = c\hat{\mathbf{p}}\gamma_0\,, \tag{2.15a}$$

$$(1+\hat{K}/2mc^2)\psi_o = (\tfrac{1}{2}mc)\hat{\mathbf{p}}\psi_e\,. \tag{2.15b}$$

Assuming $|\,\hat{K}\psi\,| << 2mc^2|\,\psi\,|$, (2.15b) can be solved for ψ_0 in the form

$$\begin{aligned}\Psi_o &= \frac{1}{2mc}\,(1+\hat{K}/2mc^2)^{-1}\,\hat{\mathbf{p}}\psi_e\\ &\approx \frac{1}{2mc}\,(1-\hat{K}/2mc^2)\,\hat{\mathbf{p}}\psi_e\,.\end{aligned} \tag{2.16}$$

Substituting this into (2.15a) and using the identity

$$(\hat{K}\hat{\mathbf{p}} - \hat{\mathbf{p}}\hat{K})\psi = \hbar\Big(-\frac{e}{c}\,\partial_t\mathbf{A} - \nabla V\Big)\psi i\boldsymbol{\sigma}_3 = e\hbar\mathbf{E}\psi i\boldsymbol{\sigma}_3\,,$$

we get

$$\hat{K}\psi_e = \frac{1}{2mc}\,\hat{\mathbf{p}}^2\psi_e - \frac{1}{4m^2c^2}\,\big(\hat{\mathbf{p}}^2\hat{K}\psi_e + e\hbar\,\hat{\mathbf{p}}(\mathbf{E}\psi_e)i\boldsymbol{\sigma}_3\big)\,. \tag{2.17}$$

Using (2.13) in the first and second terms on the right-hand side of this equation and the identity

$$e\hbar\hat{\mathbf{p}}(\mathbf{E}\psi_e)i\boldsymbol{\sigma}_3 = e\hbar^2(\nabla\mathbf{E})\psi + e\hbar(\mathbf{E}\cdot\hat{\mathbf{p}} - i\mathbf{E}\times\hat{\mathbf{p}})\psi i\boldsymbol{\sigma}_3$$

in the third term, we arrive finally at the first order correction to the Pauli equation:

$$\begin{aligned}\hat{K}\psi_e = \frac{1}{2mc}\,\hat{\mathbf{p}}^2\psi_e - \frac{e\hbar}{2mc}\,\mathbf{B}\psi_e\boldsymbol{\sigma}_3 - \frac{1}{8m^3c^2}\,\hat{\mathbf{p}}^2\psi_e\\ - \frac{e\hbar^2}{4m^2c^2}\,(\nabla\cdot\mathbf{E} + i\nabla\times\mathbf{E})\psi_e - \frac{e\hbar^2}{4m^2c^2}\,\mathbf{E}\cdot\hat{\mathbf{p}}\psi_e i\boldsymbol{\sigma}_3\\ - \frac{e\hbar}{4m^2c^2}\,\mathbf{E}\times\hat{\mathbf{p}}\psi_e\boldsymbol{\sigma}_3\,.\end{aligned} \tag{2.18}$$

With (2.8) and (2.9) it is easy to show that (2.18) is equivalent to the usual matrix equation obtained in Refs. 12–14 as the first order relativistic correction to the Pauli equation. The usual physical interpretation of (2.18) proceeds by identifying $\hat{p}_k$ as a "kinetic momentum operator" and, neglecting the difference between $\hat{p}^2$ and $\hat{\mathbf{p}}^2$ given by (2.13), interpreting the first and third terms of (2.18) as the first two terms in the expansion of a relativistic "kinetic energy operator"

$$(\hat{p}^2 + m^2c^4)^{1/2} - mc^2 = \frac{\hat{p}^2}{2m} - \frac{\hat{p}^4}{8m^3c^2} + \cdots\,. \tag{2.19}$$

The second and last terms of (2.15) are interpreted as spin precession energies, with the latter (spin-orbit) term being reduced in magnitude by a factor of $\frac{1}{2}$ attributed to the Thomas precession. The $\nabla \times \mathbf{E}$ term is usually neglected in any case it does not contribute to the energy directly. With neglect of the vector potential, the $\mathbf{E} \cdot \hat{\mathbf{p}}$ term in (2.18) can be shown to contribute to the energy the negative of half the amount of the $\nabla \cdot \mathbf{E}$ term. So, for the purpose of calculating the energy, these two terms in (2.18) can be replaced by a single term $(-e\hbar^2/8m^2c^2)(\nabla \cdot \mathbf{E})\psi_e$, the so-called "Darwin term," which is usually regarded as a quantum-mechanical effect without classical interpretation. The contributions of the last four terms in (2.18) to the energy levels of hydrogen can be evaluated by perturbation theory from the hydrogen solutions to the Schr̈odinger equation; it has been (2.14) found[16] that, as one might expect, they combine to give the α^4 term in the expansion of the Sommerfeld fine structure formula.

To sum up, the usual interpretation regards (2.18) as an approximate separation of the Dirac energy into kinetic and interaction energies. Simple and natural as this interpretation appears from the operator formulation of (2.18), it is inconsistent with the identification of kinetic energy and momentum already made in the exact Dirac theory. Nothing in the Dirac theory justifies the interpretation of (2.19) as a "kinetic energy operator." Indeed, if the operator $\hat{p}_k$ is to be interpreted as the "kinetic momentum operator," then on the basis of "relativistic invariance" alone, the operator $\hat{K}$ defined by (2.11) must be regarded as the "kinetic energy operator," which is certainly inconsistent with the interpretation of (2.18) reviewed above.

Nevertheless, the "Schrödinger approximation" to be discussed later, $\hat{K}\psi_e = (2m)^{-1}\hat{p}^2\psi_e$, showing at least that the usual interpretation of $(2m)^{-1}\hat{p}^2$ as an *approximate* kinetic energy-operator is consistent with the Dirac theory.

Section 6 of Ref. 5 gave an exact derivation of the Larmor precession energy and found that it arises from the electron mass density. Of course, in a statistical theory it is possible to identify in the mass density contributions from the kinetic and interaction energies of the statistical ensemble, but this is not allowed in the conventional Dirac theory, where the best that can be done is to determine how the mass is affected by external fields. New interpretations cannot emerge from approximations to an exact theory.

It will also shown that the Thomas precession energy came from the kinetic momentum, and to identity it the kinetic momentum was separated into two parts in Eq. (6.48) of Ref. 5. Heretofore, no one has paid attention to the fact that a similar separation was mad implicitly in the derivation of (2.18). So much emphasis is laid on the correspondence between operators and observables that it is sometimes overlooked that an operator must act on a wavefunction to produce an observable quantity. This simple fact obviously implies that a change in the wavefunction while an operator is kept fixed will generally change the correspondence with observables. Exactly this kind of change was made in arriving at (2.18). The operator $\hat{p}_k$ introduced in (2.12) can indeed (with clue attention to its relation of the energy-momentum tensor) be regarded as a momentum operator when it acts on the Dirac wavefunction ψ. It follows that the operator equation $\hat{p}_k\psi = \hat{p}_k\psi_e + \hat{p}_k\psi_o$ corresponds to a

separation of the momentum into two parts. But only $\hat{p}_k\psi_e$ is associated with momentum in (2.12) and (2.18), though $\hat{p}_k\psi_o$ is negligible only in the zeroth order (Pauli) approximation; this amounts to a change in the interpretation of the theory by identifying a different quantity as momentum. As a result, the Larmor and Thomas precession energies appear (magically) as interaction terms. Of course, as long as only the total energy is being measured experimentally it does not matter what part of it is called kinetic; only the coupling with external fields is important. However, the spin and momentum are related to one another by the angular momentum conservation law, and the interpretation of one cannot be changed without affecting the other. When this is taken into account, arbitrariness in the interpretation of various contributions to the energy is eliminated.

Before (2.18) can be correctly interpreted, the relation of ψ_e to the Dirac observables must be determined. The physical meaning of the decomposition $\psi = \psi_e + \psi_o$ is revealed by the decomposition

$$\psi = \rho^{1/2} \exp{(i\beta/2)} LU \tag{2.20}$$

obtained from Eqs. (4.2) and (6.15) of Ref. 5. The spinor L can be expressed in terms of the relative velocity $\mathbf{v}$ by taking the square root of (6.1c) in Ref. 5 to obtain (e.g., Eq. (18.14) of Ref. 17)

$$L = \frac{v_0^{1/2}}{\sqrt{2}}\Big(\frac{1}{\alpha} + \frac{\alpha v}{c}\Big)$$

$$\text{where } \alpha = \frac{v_0^{1/2}}{(v_0+1)^{1/2}} \qquad \text{and} \qquad v_0 = (1-\mathbf{v}^2/c^2)^{1/2} \,. \tag{2.21}$$

Hence (2.20) can be written

$$\psi = \frac{\exp{(i\beta/2)}}{\sqrt{2}} \left(\tfrac{1}{2} + \frac{\alpha\mathbf{v}}{c}\right) \rho_0^{1/2} U \,, \tag{2.22}$$

where of course $\rho_0 = \rho v_0$. The separation of (2.22) into even and odd parts is easily accomplished by noting that $\mathbf{v}$ and i are odd (i.e., $\gamma_0 \mathbf{v} \gamma_0 = -\mathbf{v}$ and $\gamma_0 i \gamma_0 = -i$), while U is even (i.e., $\gamma_0 U \gamma_0 = U$); hence

$$\psi_e = \frac{1}{\sqrt{2}} \Big(\frac{\cos\,\frac{1}{2}\beta}{\alpha} + i\frac{\mathbf{v}\alpha}{c} \,\sin\,\tfrac{1}{2}\beta \Big) \rho_0^{1/2} U \,, \tag{2.23a}$$

$$\psi_o = \frac{1}{\sqrt{2}} \Big(\frac{i\sin\,\frac{1}{2}\beta}{\alpha} + \frac{\mathbf{v}\alpha}{c} \,\cos\,\tfrac{1}{2}\beta \Big) \rho_0^{1/2} U \,. \tag{2.23b}$$

For $|\,\mathbf{v}/c\,| << 1$, $\alpha \approx (2)^{-1/2}$ in which case it is clear from (2.23) that $|\,\psi_o\,| << |\,\psi_e\,|$ only if β is simultaneously small (modulo π of course). Thus β small is a prerequisite for the Pauli equation to obtain as the N. R. limit of the Dirac equation. We do not attempt to explain this fact here, we merely record it as another clue to the physical interpretation of the mysterious parameter β, and we note that (2.23) shows exactly how the separation of ψ into even and odd parts depends on $\mathbf{v}/c$, a fact which we now exploit.

Expanding (2.23a,b) in β and $|\,\mathbf{v}\,|/c$ and keeping terms of first order only in both quantities, we get

$$\psi_e \approx \rho^{1/2} U \approx \rho_0^{1/2} U \equiv \chi \,, \tag{2.24a}$$

$$\psi_o \approx \tfrac{1}{2}(i\beta + \mathbf{v}/c)\,\chi \,. \tag{2.24b}$$

Substituting these in (2.7a) and multiplying on the right by $2\widetilde{\chi}$, we obtain

$$\Big(i\beta + \frac{\mathbf{v}}{c}\Big)\rho = -\frac{1}{2mc}\Big\{\hbar\nabla\chi i\boldsymbol{\sigma}_3 + \frac{e}{c}\,\mathbf{A}\chi\Big\}\widetilde{\chi} = \frac{1}{2mc}(\hat{\mathbf{p}}\chi)\widetilde{\chi} \,. \tag{2.25a}$$

Separating this into relative vector and pseudovector parts, we get

$$\mathbf{v} = -\frac{\hbar}{m\,\rho}\,[\nabla\chi i\boldsymbol{\sigma}_3\widetilde{\chi}]_{(1)} - \frac{e}{c}\,\mathbf{A} \tag{2.25b}$$

and

$$\beta = -\frac{\hbar}{m\,c\,\rho}\,[\nabla\chi\boldsymbol{\sigma}_3\widetilde{\chi}]_{(0)} = -\frac{1}{m\,c}\,\frac{\hbar}{2}\,\nabla\cdot(\chi\boldsymbol{\sigma}_3\widetilde{\chi}) = -\frac{1}{m\,c\rho}\,\nabla\cdot(\rho\mathbf{s}) \,, \tag{2.25c}$$

where we have used (6.16) and (7.3a) of Ref. 5 to identify the relation of the spin to the Pauli wavefunction χ as

$$\rho\mathbf{s} = \tfrac{1}{2}\hbar\rho U\boldsymbol{\sigma}_3\widetilde{U} = \tfrac{1}{2}\hbar\chi\boldsymbol{\sigma}_3\widetilde{\chi} \,. \tag{2.26}$$

The result (2.25c) is just an approximate derivation of (7.3a) in Ref. 5, which has already been shown to hold under more general assumptions. Equation (2.25b) is the correct expression for the electron velocity in terms of the Pauli wavefunction. This reveals the physical significance of Eq. (2.7a) in the Pauli theory.

We are now in a position to examine the physical significance of the spin-electric coupling terms in (2.18) by expressing them in terms of local observables. The last three terms in (2.18) are equivalent to the single term in (2.17)

$$\begin{aligned} -\frac{e\hbar}{4m^2c^2}\,\hat{\mathbf{p}}(\mathbf{E}\psi_e)i\boldsymbol{\sigma}_3 &= -\frac{e\hbar^2}{4m^2c^2}\,\nabla(\mathbf{E}\psi_e) \\ &= -\frac{e\hbar^2}{4m^2c^2}\,[\,(\nabla\mathbf{E})\psi_e + 2\mathbf{E}\cdot\nabla\psi_e - \mathbf{E}\nabla\psi_e\,] \,, \end{aligned} \tag{2.27}$$

which we have reexpressed by using (2.12a) and neglecting the vector potential $\mathbf{A}$. In the present approximation, we can replace ψ_e in (2.27) by the Pauli wavefunction $\chi = \rho^{1/2}U$, whereupon, after multiplication on the right by $\widetilde{\chi}$, (2.27) becomes

$$-\frac{e\hbar}{4m^2c^2}\,(\hat{\mathbf{p}}\mathbf{E}\chi)i\boldsymbol{\sigma}_3\widetilde{\chi} = -\frac{e\hbar^2}{4m^2c^2}\Big(\rho\nabla\mathbf{E} + \mathbf{E}\cdot\nabla\rho + 2(\mathbf{E}\cdot\nabla U)\widetilde{U}\Big) + \frac{e}{2mc}\,\mathbf{E}\Big(\frac{\hbar^2}{2mc}(\nabla\chi)\widetilde{\chi}\Big)\,. \tag{2.28}$$

But if (2.25a) is multiplied by $i\mathbf{s}$ and (2.26) is used, one finds, neglecting $\mathbf{A}$,

$$\frac{\hbar^2}{2mc}\,(\nabla\chi)\widetilde{\chi} = \Big(i\,\frac{\mathbf{v}}{c} - \beta\Big)\,\rho\mathbf{s}\,. \tag{2.29}$$

Noting that $\mathbf{s}^2 = \frac{1}{4}\hbar^2$, that $(\mathbf{E}\cdot\nabla U)\widetilde{U}$ is a bivector, and that $i\mathbf{E}\mathbf{v}\mathbf{s} = \mathbf{E}\cdot\mathbf{v}i\mathbf{s} - (\mathbf{E}\times\mathbf{v})\mathbf{s}$, we obtain, on substituting (2.29) and (2.28) and taking the scalar part,

$$-\frac{e\hbar}{4m^2c^2}\,[(\hat{\mathbf{p}}\mathbf{E}\chi)i\boldsymbol{\sigma}_3\widetilde{\chi}]_{(0)} = -\frac{e\mathbf{s}^2}{m^2c^2}\,\nabla\mathbf{E}\cdot(\rho\mathbf{E}) - \frac{e}{2mc^2}\,\rho\mathbf{s}\cdot(\mathbf{E}\times\mathbf{v}) - \frac{e}{2mc}\,\mathbf{s}\cdot\mathbf{E}\rho\beta\,. \tag{2.30}$$

Of course, the perfect divergence $\nabla\cdot(\rho\mathbf{E})$ in (2.30) has a vanishing contribution to the total energy so, with $\sin\beta = \beta$, (2.30) is seen to differ from the expression (7.14) of Ref. 5 for the spin-electric energy density essentially by a factor $\frac{1}{2}$ in the last term. We shall not look into the reason for this discrepancy, because it does not affect matters of interpretation which concern us now.

The importance of the last term in (2.30) depends on the magnitude of β relative to $|\,\mathbf{v}/c\,|$, and that can be determined only by computation from the solution to the wave equation. The hydrogen atom solutions to the Pauli equation gives $\mathbf{s}$ constant, in which case (2.25c) gives $\beta = -(mc)^{-1}\mathbf{s}\cdot\nabla\ln\rho$; so if ρ is a sufficiently slowly varying function of position, we have $\beta << |\,\mathbf{v}/c\,|$, and (2.30) is equivalent to (7.14) and (7.17) in Ref. 5, giving us in this approximation

$$\rho E_{\mathrm{SE}} = -\frac{e}{2mc^2}\,\rho\mathbf{s}\cdot(\mathbf{E}\times\mathbf{v}) = -\frac{e}{2m^2c^2}\,\{\rho\mathbf{s}\cdot(\mathbf{E}\times\mathbf{p}) + \mathbf{s}^2\big(\rho\nabla\cdot\mathbf{E} - \nabla\cdot(\rho\mathbf{E})\big)\}\,. \tag{2.31}$$

This shows that the hydrogen spin-electric energy given by (2.18) is identical to the one arrived at by Thomas from purely classical considerations (see Ref. 1). This fact is disguised in (2.18) by the use of operators and the failure to distinguish between velocity and momentum.

The mysterious Darwin term is completely explained by (2.31). Since $\nabla\cdot\mathbf{E}$ is proportional to a delta function vanishing everywhere except at the origin and ρ is nonvanishing at the origin only for s-states, the Darwin term contributes only to s-states. But $\mathbf{p} = 0$ for s-states, so (2.31) shows that the Darwin term gives the entire spin-orbit coupling for s-states. Even though the momentum density $\rho\mathbf{p}$ vanishes for s-states, spin-orbit coupling is possible because the charge current $e\rho\mathbf{v} = em^{-1}\nabla\times\rho\mathbf{s}$ is finite. The numerical coefficient of the Darwin term is notable: as (2.31) shows a factor $\hbar^2/8$ arises from the "Thomas factor" $\frac{1}{2}$ and the spin $\mathbf{s}^2 = \frac{1}{4}\hbar^2$.

For other than s-states the distinction between velocity and momentum is not so important, being responsible only for a term $-\frac{1}{2}em^{-2}c^{-2}\mathbf{E}\cdot\mathbf{s}\mathbf{s}\cdot\nabla\rho$ which was neglected in relating (7.17) of Ref. 5 to (2.31) here. This is to say that the magnetization current makes an important contribution to the energy only for s-states, where it is the entire current.

The identification of Thomas precession in the Dirac theory is justified in current textbooks solely by noting the factor $\frac{1}{2}$ in the spin orbit term of (2.18) which remains after associating a factor $\frac{1}{2}\hbar$ with the spin. The identification of the "Thomas factor" is correct, as we have shown by the same general argument as Thomas. based primarily on the fact that the proper spin S is always orthogonal to the proper particle velocity v. But its appearance in (2.18) is rather fortuitous, because the separation of the Dirac momentum into $\hat{p}_k\psi_k + p_k\psi_0$ which makes the Thomas precession explicit in (2.18) is equivalent to the exact separation made in (6.48) of Ref. 5 only to first order. It must be remembered that, as cautioned in Ref. 5, it is possible to talk of the Thomas precession only when $\Sigma = Ui\boldsymbol{\sigma}_3\widetilde{U}$ is taken to be the spin.

3. Observables in the Pauli-Schrödinger Theory

In the introduction we pointed out that the interpretation of the Schrödinger theory is a theory of an electron without spin is inconsistent with the view that it is an approximation to the Dirac theory. *Consistency requires that the Schrödinger theory be regarded as describing an electron in an eigenstate of spin.* Here the term "eigenstate" can be taken in the usual sense. But we think that the sense suggested in Sec. 5 of Ref. 5 is more revealing. Accordingly, *we say that an electron is in an eigenstate of the spin if and only if the local spin vector* $\mathbf{s} = U\boldsymbol{\sigma}_3\widetilde{U}$ is uniform, i.e., constant in time and homogeneous in space.

To emphasize the fact that the Schrödinger theory is identical to the Pauli theory for an electron in an eigenstate of spin, we speak of the "Pauli-Schrödinger (P-S) theory." The P-S theory has already been discussed in Ref. 6, and everything mentioned there is consistent with the Dirac theory. But the derivation of the P-S theory from the Dirac theory which has been carried out in the preceding sections reveals some important features of the P-S theory which were not mentioned in Ref. 6. This has particularly significant consequences for the interpretation of the Schrödinger theory.

Let us summarize the *assumptions* of the P-S theory which we have *derived* from the Dirac theory in Section 2. The P-S wavefunction χ can be written in the form

$$\chi = \rho\tfrac{1}{2}U\,, \tag{3.1a}$$

where

$$UU^\dagger = 1 \tag{3.1b}$$

$$U^\dagger \equiv \gamma_0\widetilde{U}\gamma_0 = \widetilde{U} \tag{3.1b}$$

and

$$\rho = \chi^\dagger\chi = \chi\chi^\dagger \tag{3.2a}$$

is a scalar to be interpreted as the probability density. The wavefunction χ is determined by assuming that it is a solution of the Pauli equation (2.7b) or (2.13) or emphasized in Ref. 6, the Pauli equation reduces to the Schrödinger equation when the magnetic field is sufficiently small. Besides the probability density (3.2a), the fundamental observables of the P-S theory are the energy density ρE, the momentum density $\rho\mathbf{p}$, the spin density $\rho\mathbf{s}$, the charge current $\mathbf{j} = e\rho\mathbf{v}$; they are expressible in terms of the wavefunction by the equations

$$\rho E = \hbar(\partial_t\chi i\boldsymbol{\sigma}_3\widetilde{\chi})_{(0)} = \hbar(\partial_t U i\boldsymbol{\sigma}_3 U^\dagger)_{(0)}U\,, \tag{3.2b}$$

$$\rho\mathbf{s} = \frac{\hbar}{2}\,\chi\boldsymbol{\sigma}_3\widetilde{\chi} = \frac{\hbar}{2}\,\rho U\boldsymbol{\sigma}_3 U^\dagger\,, \tag{3.2c}$$

$$(\hat{p}_k\chi)\chi^\dagger \equiv -\Big\{\hbar\partial_k\chi i\boldsymbol{\sigma}_3 + \frac{e}{c}\,A_k\chi\Big\}\chi^\dagger = \rho p_k - i\partial_k(\rho\mathbf{s})\,, \tag{3.2d}$$

$$m^{-1}(\hat{\mathbf{p}}\chi)\chi^\dagger \equiv -m^{-1}\Big\{\hbar\nabla\chi i\boldsymbol{\sigma}_3 + \frac{e}{c}\,\mathbf{A}\chi\Big\}\chi^\dagger = \rho(\mathbf{v} + ic\beta)\,. \tag{3.2e}$$

This completes the list of assumptions derived for the P-S theory.

For purposes of comparison, we use (2.8) and (2.9) to express the observables (3.2) in the usual matrix; notation. Writing $\Psi = \chi u_1$, as in (2.9a), for the matrix wavefunction, introducing $\hat{p}_k\Psi \equiv -(i'\hbar\partial_k + (e/c)A_k)\Psi$ and using "Re" to denote "real part," we get

$$\rho = \chi^\dagger\chi = (\chi^\dagger\chi)_{(0)} = \Psi^\dagger\Psi\,, \tag{3.3a}$$

$$\rho E = \hbar(\chi^\dagger\partial_t\chi i\boldsymbol{\sigma}_3)_{(0)} = \mathrm{Re}\{i'\hbar\Psi^\dagger\partial_t\Psi\}\,, \tag{3.3b}$$

$$\rho s_k = \rho\mathbf{s}\cdot\boldsymbol{\sigma}_k = \tfrac{1}{2}\hbar(\boldsymbol{\sigma}_k\chi\boldsymbol{\sigma}_3\chi^\dagger)_{(0)} = \tfrac{1}{2}\hbar\Psi^\dagger\boldsymbol{\sigma}_k\Psi\,, \tag{3.3c}$$

$$\rho p_k = \rho(U^\dagger\hat{p}_k U)_{(0)} = (\chi^\dagger\hat{p}_k\chi)_{(0)} = \mathrm{Re}\{\Psi^\dagger\hat{p}_k\Psi\}\,, \tag{3.3d}$$

$$\rho v_k = \rho\mathbf{v}\cdot\boldsymbol{\sigma}_k = (\chi^\dagger\hat{p}_k\chi)_{(0)} = m^{-1}\mathrm{Re}\{\Psi^\dagger\boldsymbol{\sigma}_k\boldsymbol{\sigma}_j\hat{p}_j\Psi\}\,. \tag{3.3e}$$

Equations (3.3) can be used instead of (3.2) to relate observables to the wavefunction, but (3.2) is easier to work with. For example using $m^{-1}\hat{\mathbf{p}} = m^{-1}\boldsymbol{\sigma}_k\hat{p}_k$, we get immediately from (3.2d) and (3.2e)

$$\rho\mathbf{p} - i\nabla(\rho\mathbf{s}) = m\rho(\mathbf{v} + ic\beta)\,. \tag{3.4a}$$

Since $\nabla(\rho\mathbf{s}) = \nabla\cdot(\rho\mathbf{s}) + i\nabla\times(\rho\mathbf{s})$ the vector part of (3.4a) is

$$m\rho\mathbf{v} = \rho\mathbf{p} + \nabla\times(\rho\mathbf{s})\,, \tag{3.4b}$$

while the pseudovector part gives

$$mc\rho\beta = -\nabla\cdot(\rho\mathbf{s})\,. \tag{3.4c}$$

In Sec. 7 of Ref. 5 the fundamental relation (3.4b) was derived from a "constitutive equation" determined by the Dirac equation. In Sec. 2 we saw that a decoupling of (3.4b) from the wave equation was brought about by the separation of the Dirac wavefunction into large and small components. So (3.4b) must be introduced into the P-S theory as an assumption independent of the P-S wave equation. It has already been pointed out in Sec. 7 of Ref. 5 that (3.4b) expresses the separation of the total charge current $j = e\rho\mathbf{v}$ into a convection current $em^{-1}\rho\mathbf{p}$ and a magnetization current $\nabla\times(em^{-1}\rho\mathbf{s})$ written as Eq. (1.2) in the introduction. The physical significance of (3.4c) is not is obvious as that of (3.4b). Indeed, since the function β is given in terms of ρ and $\mathbf{s}$ by (3.4c), it need not be introduced into the theory at all. However, we have already observed the peculiar role of β in the Dirac theory, so it should be interesting to see how it enters the P-S theory.

From the set of basic observables (3.2) other observables can be constructed. Chief among these are the "orbital angular momentum density"

$$\rho\mathbf{L} \equiv \mathbf{x}\times(\rho\mathbf{p}) = \rho\mathbf{x}\times\mathbf{p}\,. \tag{3.5}$$

and the "total angular momentum density"

$$\rho J \equiv \rho(\mathbf{L} + \mathbf{s})\,. \tag{3.6}$$

Since, according to (3.4b), $\mathbf{p} \neq m\mathbf{v}$, the moment of charge differs from the moment of momentum. To express this difference we introduce the orbital momentum density

$$\rho\mathbf{L}^* \equiv \mathbf{x}\times(\rho m\mathbf{v}) = m\rho\mathbf{x}\times\mathbf{v}\,. \tag{3.7}$$

By virtue of (3.4b), the relation between the two orbital moment densities can be put in the form

$$\rho\mathbf{L}^* = \rho\mathbf{L} + 2\rho\mathbf{s} + \partial_k\{\rho\mathbf{x}\times(\boldsymbol{\sigma}_k\times\mathbf{s})\}\,. \tag{3.8}$$

Other important observables appear when we use the Pauli equation to derive equations of motion for the local observables. With $\mathbf{p}$ and $\mathbf{v}$ related to the wave equation by (3.2d,e), it is easy to show that the Pauli equation implies the conservation laws

$$\partial_t\rho + \nabla\cdot(m^{-1}\rho\mathbf{p}) = 0\,, \tag{3.9}$$

$$\partial_t\rho + \nabla\cdot(\rho\mathbf{v}) = 0\,. \tag{3.10}$$

Indeed, either of these equations follows from the other by virtue of (3.4b). Equation (3.9) implies the existence of momentum streamlines. We can always write the equation for momentum conservation on such a streamline in the general form

$$\rho D_t\mathbf{p} \equiv \rho\big(\partial_t + (\mathbf{p}/m)\cdot\nabla\big)\mathbf{p} = \rho\mathbf{f}' - \partial_k\mathbf{N}'_k\,. \tag{3.11a}$$

This introduces the *force density* $\rho\mathbf{f}'$ and the stress $\mathbf{N}'_k$ on a volume element moving with the streamline as local observables. In Ref. 6. the Pauli equation was shown to lead to the specific expressions

$$\mathbf{f}' = e\{\mathbf{E} + (\mathbf{p}/mc)\times\mathbf{B}\} + (e/mc)\boldsymbol{\sigma}_k\mathbf{s}\cdot\partial_k\mathbf{B}\,, \tag{3.11b}$$

$$\mathbf{N}'_k = -M^{-1}\rho\boldsymbol{\sigma}_j\mathbf{s}\cdot(\partial_j\partial_k\mathbf{s} + \mathbf{s}\partial_j\partial_k\ln\rho)\,, \tag{3.11c}$$

for $\mathbf{f}'$ and $\mathbf{N}'_k$ in terms of the basic local observables (3.2).

On the other hand, Eq. (3.10) rather than (3.9) is the charge conservation equation, and momentum conservation along a *charge* (or *velocity*) streamline has the general form

$$\rho\, d_t\mathbf{p} \equiv \rho(\partial_t + \mathbf{v}\cdot\nabla)\mathbf{p} = \rho\mathbf{f} - \partial_k\mathbf{N}_k\,. \tag{3.12a}$$

By the method of Ref. 6 it can be shown that the Pauli equation leads to the expressions

$$\mathbf{f} = e\{\mathbf{E} + (\mathbf{v}/c)\times\mathbf{B}\}\,, \tag{3.12b}$$

$$\mathbf{N}_k = \rho c\boldsymbol{\sigma}_k\cdot\mathbf{s}\nabla\beta + \boldsymbol{\sigma}_j\mathbf{s}\cdot(\boldsymbol{\sigma}_k\times\partial_j\mathbf{v})\,. \tag{3.12c}$$

We will not carry out the derivation here, because in Sec. 4 the same results will be obtained from the hydrodynamic formulation of the Dirac theory. The corresponding conservation laws for angular momentum will also be obtained. Our object here is only to compare the conservation law (3.11) and (3.12) on the momentum and charge streamlines.

The structure of the stress terms (3.11c) and (3.12b) appears to be difficult to understand. But (3.12b) shows that the force on a charge streamline is exactly the classical “Lorentz force”; this by itself is nearly sufficient to show that the electromagnetic interaction in quantum theory is the same as in classical theory. On the other hand, (3.11b) shows that the force on a momentum streamline consists of a “Lorentz force” supplemented by a “Stern-Gerlach force.” In the light of (3.12b), we conclude that the “Stern-Gerlach force” arises from the circulation of charge relative to the momentum streamlines. The same general conclusion was reached in Ref. 4 by an examination of the hydrodynamic formulation Dirac theory. But it should be recalled that the analysis there is complicated by the fact that there are no momentum streamlines in the exact Dirac theory though there are streamlines generated by the Gordon current.

We have summarized how observables are brought into the P-S theory. A basic set such as (3.2) must be defined in terms of the wavefunction, while the remaining observables arise when the wave equation is used to construct the conservation laws. In this connection it is well to recall that, as shown in Ref. 6, even the spin does not have to be introduced by definition in the Dirac theory,, because it is determined by the conservation laws. The conservation laws are equations of motion for the observables, and as we have seen with (3.11) and (3.12) they have properties which can be interpreted physically. In this way the physical consequences of the wave equation are revealed.

So far we have discussed only local observables in the P-S theory. The “average” (or “global”) observables obtained by averaging local observables over spaces are easier to study experimentally, so we now ascertain some of their properties.

The average momentum has the classical relation to the average velocity

$$\langle\,\mathbf{p}\,\rangle \equiv \int d^3x\rho\mathbf{p} = m\langle\,\mathbf{v}\,\rangle\,. \tag{3.13}$$

This follows from (3.4b) since the contribution of the “spin current” vanishes by Gauss’s theorem. However, by integrating (3.8) we find that the angular momentum differs from the average moment of charge according to the formula

$$\langle\,\mathbf{L}^*\,\rangle = \langle\,\mathbf{L}\,\rangle + 2\langle\,\mathbf{s}\,\rangle\,. \tag{3.14}$$

Multiplying the Pauli equation taking the scalar part and integrating we get the expression for the average kinetic energy

$$\langle\,\hat{K}\,\rangle = \left\langle\,\frac{\hat{p}^2}{2m}\,\right\rangle - \frac{e}{mc}\langle\,\mathbf{B}\cdot\mathbf{s}\,\rangle = \langle\,E\,\rangle - \langle\,V\,\rangle\,, \tag{3.15}$$

where, in particular,

$$\langle\, \hat{p}^2 \,\rangle = \int d^3x (\chi^\dagger \hat{p}^2 \chi)_{(0)}\,, \tag{3.16}$$

$$\langle\, \mathbf{B}\cdot\mathbf{s} \,\rangle = \int d^3x \rho \mathbf{B}\cdot\mathbf{s}\,. \tag{3.17}$$

The "operator expectation" (3.16) can be reexpressed in terms of the velocity, with the striking result

$$\langle\, \hat{K} \,\rangle = \langle\, \tfrac{1}{2} m\mathbf{v}^2 \,\rangle + \langle\, \tfrac{1}{2} mc^2 \beta^2 \,\rangle\,. \tag{3.18}$$

In Sec. 4 this result will be obtained from an exact relation holding, in the Dirac theory

According to (2.15) the operator $(2m)^{-1}\hat{p}^2$ can be interpreted as the *kinetic energy* operator only if $\langle\, \mathbf{B}\cdot\mathbf{s} \,\rangle = 0$, that is, in the Schr̆odinger approximation to the Pauli theory. The expression $\langle\, \hat{p}^2 \,\rangle$ is unsatisfactory from our point of view, because it has not been expressed in terms of the basic local observables in (2.2). We can easily reexpress it in terms of the momentum (2.3d) by using Eq. (4.11) of Ref. 6, which shows us that

$$(\chi^\dagger \hat{p}^2 \chi)_0 = \rho \big\{ \mathbf{p}^2 - \mathbf{s}^2 [2\nabla^2 \ln \rho + (\nabla \ln \rho)^2] - \mathbf{s}\cdot(\nabla^2 \mathbf{s}) \big\}\,. \tag{3.19}$$

Hence

$$\langle\, \hat{p}^2 \,\rangle = \langle\, \mathbf{p}^2 \,\rangle - \langle\, \mathbf{s}^2 [2\nabla^2 \ln \rho + (\nabla \ln \rho)^2] - \mathbf{s}\cdot(\nabla^2 \mathbf{s}) \,\rangle\,. \tag{3.20}$$

Equation (3.20) shows, at least, that $\langle\, p^2/2m \,\rangle$ does not give the entire contribution to the kinetic energy, so, on the basis of general principles of continuum mechanics, one is tempted to interpret the spin and density terms in (3.20) as a kind of heat energy associated with the local angular momentum flux, but it is difficult to account for the specific form of the terms on the basis of this idea.

Obviously, the kinetic energy is much more simply expressed in terms of the velocity by (3.18) than in terms of the momentum. Notice that, in contrast to (3.15), (3.18) displays no explicit interaction with the magnetic field. This is entirely in accordance with the idea that the spin arises from a circulation of charge. But the problem remains to understand the β^2 term in (3.18). We shall comment on this later.

A virial theorem for the P-S theory is easily derived with the help of (3.12) but not by using (3.11)!]. Since, in Sec. 4 we will obtain a more general theorem from the Dirac theory by a similar method, we do not give the derivation. Looking ahead, we merely note that from Eqs. (4.20)–(4.23) we get, recalling that the rest energy is omitted from the Pauli theory, the general virial theorem for stationary states

$$\langle\, E \,\rangle = -\langle\, \tfrac{1}{2} m\mathbf{v} \,\rangle - \langle\, \tfrac{1}{2} mc^2 \beta^2 \,\rangle + \langle\, V \,\rangle - e\langle\, \mathbf{x}\cdot(\mathbf{E} + c^{-1}\mathbf{x}\times\mathbf{B}) \,\rangle\,, \tag{3.21}$$

which for a Coulomb field alone reduces to

$$\langle\, E \,\rangle = -\langle\, \tfrac{1}{2} m\mathbf{v} \,\rangle - \tfrac{1}{2} mc^2 \langle\, \beta^2 \,\rangle\,. \tag{3.22}$$

Equating this to $\langle\, E \,\rangle + \langle\, K \,\rangle + \langle\, V \,\rangle$ and using (3.18), we get the alternative form of the virial theorem

$$\langle\, \mathbf{v}^2 \,\rangle + mc^2 \langle\, \beta^2 \,\rangle + \langle\, V \,\rangle = 0\,. \tag{3.23}$$

Let us consider some implications of the above relations among observables for the interpretation of the Schrödinger solution of the hydrogen atom. As has been explained, we get the Schrödinger theory *with spin* by taking $\mathbf{s}$ constant. This determines a preferred direction in the theory, so it is convenient to write

$$\mathbf{s} = \tfrac{1}{2}\hbar \boldsymbol{\sigma}_3\,. \tag{3.24}$$

Obviously,

$$\langle\, \mathbf{s} \,\rangle = \mathbf{s} = \tfrac{1}{2}\hbar \boldsymbol{\sigma}_3 \tag{3.25}$$

for any solution of the Schrödinger equation. From the Schrödinger wavefunctions for the hydrogen atom we get

$$\langle\, \mathbf{L} \,\rangle = m\hbar \boldsymbol{\sigma}_3 = 2m\mathbf{s}\,, \tag{3.26}$$

where m is the magnetic quantum number. This is identical to the usual result

$$\langle\, \hat{L}_3 \,\rangle = m\hbar\,, \qquad \langle\, \hat{L}_1 \,\rangle = 0 = \langle\, \hat{L}_2 \,\rangle\,,$$

where the $\hat{L}_k$ the usual angular momentum operators, that is,

$$\langle\, \hat{L}_k \,\rangle = \boldsymbol{\sigma}_k \cdot \langle\, \mathbf{L} \,\rangle = \langle\, \boldsymbol{\sigma}_k \cdot \mathbf{L} \,\rangle\,.$$

The total angular momentum is clearly

$$\langle\, \mathbf{J} \,\rangle = (m + \tfrac{1}{2})\hbar\boldsymbol{\sigma}_3 = (2m+1)\mathbf{s}\,. \tag{3.27}$$

The spin is coupled to other observables by Eq. (3.46), which, in the Schrödinger theory can be written

$$\mathbf{m}\rho v = \rho\mathbf{p} - \mathbf{s} \times \nabla\rho\,. \tag{3.28}$$

For s-states the Schrödinger wavefunctions imply $\mathbf{p} = 0$, but (3.28) shows that $\mathbf{v} \neq 0$. In fact, since ρ is spherically symmetric $\nabla\rho$ is directed radially, so (3.28) implies that the charge streamlines circulate about the spin axis. Thus, the *the charge distribution of an s-state is not static* as it is usually supposed to be when $(e/m)\mathbf{p}$ is wrongly assumed to describe the charge flow. It must be emphasized that this conclusion is solely a consequence of requiring that the Schrödinger theory be a consistent approximation to the Dirac theory. Of course, we have already seen that the charge current in the s-state, gives the "Darwin spin-orbit energy" when the first order relativistic correction is included.

The correspondence between Schrödinger and Bohr theories of hydrogen is significantly improved by corresponding $\mathbf{v}$ rather than $\mathbf{p}/m$ with the velocity of the electron in the Bohr theory. All observables can be expressed in terms of $\mathbf{v}$ rather than $\mathbf{p}/m$ by using (3.28). The angular momentum to compare with the Bohr theory is therefore $\mathbf{L}^*$ defined by (3.7) rather than $\mathbf{L}$. Indeed, substitution of (3.25), (3.26) into (3.14) yields

$$\langle\, L^* \,\rangle = (m+1)\hbar\boldsymbol{\sigma}_3\,. \tag{3.29}$$

As in the Bohr theory, (3.29) associates a finite angular momentum with the s-states. Since spin is separately conserved it appears that we can just omit it from the angular momentum balance. But a deeper analysis may show that the half-integral values of the "azimutal quantum number" in the Schrödinger theory are related to the spin.

It is tempting to suppose that the Schrödinger theory describes some sense a statistical ensemble of Bohr-like orbits. The expression (3.18) for the kinetic energy seems to be very close to what one might expect in such a case. Consider the strange quantity β which appears there. From (3.4c) we have

$$\beta = -\frac{1}{mc}\,\nabla \cdot (\rho\mathbf{s}) = -\frac{\hbar}{2mc}\,\boldsymbol{\sigma}_3 \cdot \ln\,\rho\,, \tag{3.30}$$

so in the Schrödinger theory (3.12) can be written

$$\langle\, \hat{K} \,\rangle = \tfrac{1}{2}m\langle\, \mathbf{v}^2 \,\rangle + \frac{\mathbf{s}^2}{2m}\left\langle \left(\frac{\partial \ln\,\rho}{\partial z}\right)^2 \right\rangle\,. \tag{3.31}$$

Thus the contribution of P to the energy is determined by by the derivative of the density along the "axis of quantization." But the Bohr orbits are confined to a plane, while the Schr̈odinger $\mathbf{v}$-orbits are distributed throughout space, though they too circulate about a preferred axis. Perhaps the last term in (3.31) is only needed to compensate for this difference or perhaps it represents the entire contribution of the spin to the energy. Perhaps the biggest difference between the Bohr and Schrödinger theories is that the latter contains spin. No doubt the last word on the subject has not been spoken.

4. Hydrodynamic Equations for Relative Observables

In Sec. 6 of Ref. 5 the fundamental *relative* local observables, velocity, spin, energy, and momentum were introduced and the basic constitutive relations among them were ascertained and discussed. This section derives the *relative* equations of motion for these quantities. The resulting equations are exact but probably too complicated to be of practical interest. We use them only to derive a virial theorem and to find the corresponding hydrodynamic equations of the Pauli-Schrödinger theory in the nonrelativistic limit. However, it may be possible to use these equations to compute relativistic dynamical corrections to the Pauli theory, a problem which is difficult to handle with a wave equation.

Let us recall the dynamical conservation laws of the Dirac theory. According to Eqs. (2.26) and (3.22) of Ref. 4, the energy-momentum conservation law can be written

$$\rho\, d_\tau p = \frac{e}{c}\,\rho F\cdot v - \partial_\mu N^\mu\,, \tag{4.1a}$$

where

$$N(\gamma^\mu) = N^\mu = \gamma^\nu \rho(v\wedge\gamma^\mu)\cdot W = -\rho s^\mu \Box\beta + \gamma^\nu\rho[v\gamma^\mu\partial_\nu S]_{(0)}\,. \tag{4.1b}$$

Note that $N(v) = v_\mu N^\mu = 0$, =0, so $N(n)$ can be identified as the proper stress tensor, describing the flux of momentum through a hypersurface with normal n into a particle streamline. Equation (4.1a) says that the Lorentz force is the only body force on the electron; the "Stern-Gerlach" force does not appear explicitly in (4.1a) because it is not a body force, rather it has been shown in 4 to arise from the term $\partial_\mu N^\mu$, so it expresses the influence of the external field on the local momentum flux.

According to Eqs. (2 34) and (2.33) of Ref. 4, angular momentum conservation in the Dirac theory can be expressed by the equation

$$\rho\, d_\tau S = -\partial_\mu M^\mu + \rho v\wedge p + \gamma_\mu N^\mu \tag{4.2a}$$

where

$$M(\gamma^\mu) \equiv M^\mu = \rho S\cdot\gamma^\mu v = \rho\tfrac{1}{2}[S,\gamma^\mu\wedge v]\,. \tag{4.2b}$$

The last two terms of (4.2a) describe the coupling of the spin to the energy-momentum density a flux via the skew-symmetric part of the energy-momentum tensor. Also $M(v) = 0$ so the tensor $M(n)$ gives the flux of angular momentum in the direction n onto a streamline Thus (4.2a) says that the spin is subject to no body torques: the "Larmor term" does not appear explicitly in (4.2a), because, as shown in Ref. 4, it describes the influence of the external field on the local angular momentum flux.

The main task of this section is to reexpress the conservation laws (4.1) and (4.2) in relative form. This work completes the job of expressing the Dirac theory in terms of relative variables which was begun in Sec. 6 of Ref. 5, so familiarity with the characterization of relative variables developed here is presumed.

Consider first the conservation Eq. (4.1a) for energy-momentum. Using (6.6) and (6.7) of Ref. 5, we write

$$\rho\, d_\tau p\gamma_0 = v_0 c^{-1}\rho_0 d_t\Big(\frac{\epsilon}{c} + \mathbf{p}\Big)\,. \tag{4.3}$$

Also it is easy to show that [Eq. (2.15) of Ref. 11]

$$F\cdot v\gamma_0 = v_0[c^{-1}\mathbf{E}\cdot\mathbf{v} + c^{-1}\mathbf{v}\times\mathbf{B}]\,. \tag{4.4}$$

Hence after multiplication by γ_0, Eq. (4.1a) can be separated into the two equations

$$\rho_0\, d_t\epsilon = \rho_0\, e\,\mathbf{E}\cdot\mathbf{v} - \partial_\mu N_0^\mu\,, \tag{4.5a}$$

$$\rho_0\, d_t\mathbf{p} = \rho_0\, e\,(\mathbf{E} - c^{-1}\mathbf{v}\times\mathbf{B}) - \partial_\mu\mathbf{N}^\mu\,, \tag{4.5b}$$

where we have introduced the notation

$$N_0^\mu = c^2 N^\mu\cdot\gamma_0\,, \qquad \mathbf{N}^\mu = cN^\mu\wedge\gamma_0\,.$$

To get the flux terms expressed in terms of relative local observables, we use (4.26).

Introducing the frame of relative vectors

$$\boldsymbol{\sigma}_k \equiv \gamma_k\gamma_0 = -\gamma_0\gamma_k = \gamma_0\gamma^k \qquad (k = 1,2,3)\,,$$

and recalling (6.4b) of Ref. 5, we have

$$c^{-1}N_0^\mu = -\rho_0\, s^\mu\partial_t\beta + \rho_0\,[v\gamma^\mu\partial_t S]_{(0)}\,, \tag{4.6a}$$

$$c^{-1}\mathbf{N}^{\mu} = \rho_0\, s^{\mu}\nabla\beta - \boldsymbol{\sigma}_j\rho_0\,[v\gamma^{\mu}\partial_j S]_{(0)}\,, \tag{4.6b}$$

and, recalling (6.1) and (6.12a) of Ref. 5, we have

$$\begin{aligned}
[v\gamma^0\partial_\nu S]_{(0)} &= \frac{v_0}{c}\,[\mathbf{v}\partial_\nu(\mathbf{s}_1+S_2)]_{(0)} = \frac{v_0}{c}\,\mathbf{v}\cdot\partial_\nu\mathbf{s}_1\\
[v\gamma^k\partial_\nu S]_{(0)} &= [v\gamma_0\gamma_0\gamma^k\partial_\nu S]_{(0)}\\
&= v_0[(1+\mathbf{v}/c)\boldsymbol{\sigma}_k\partial_\nu(\mathbf{s}_2+S_2)]_{(0)}\\
&= v_0\left\{\partial_\nu(\boldsymbol{\sigma}_k\cdot\mathbf{s}_1) + c^{-1}(\mathbf{v}\wedge\boldsymbol{\sigma}_k)\cdot\partial_\nu S_2\right\}.
\end{aligned}$$

Hence,

$$N_0^0 = -c\rho\, s_0\partial_t\beta + \rho_0\mathbf{v}\cdot\partial_t\mathbf{s}_1\,, \tag{4.7a}$$

$$N_0^k \equiv N_{0k} = -c\rho\, s^k\partial_t\beta + \rho_0\big\{c\partial_t\mathbf{s}_1\cdot\boldsymbol{\sigma}_k + (\mathbf{v}\wedge\boldsymbol{\sigma}_k)\cdot\partial_t S_2\big\}\,, \tag{4.7b}$$

$$\mathbf{N}^0 = c\rho\, s_0\nabla\beta - \rho_0\boldsymbol{\sigma}_j\mathbf{v}\cdot\partial_j\mathbf{s}_1\,, \tag{4.7c}$$

$$\mathbf{N}^k \equiv \mathbf{N}_k = c\rho\, s^k\nabla\beta - \rho_0\big\{c\nabla\mathbf{s}_1\cdot\boldsymbol{\sigma}_k + \boldsymbol{\sigma}_j(\mathbf{v}\wedge\boldsymbol{\sigma}_k)\cdot\partial_j S_2\big\}\,. \tag{4.7d}$$

It will be noted that these flux term, are expressed in terms of several quantities which are not independent of one another. By (6.12), (6.18), (6.19) of Ref. 5 and (2.18) of Ref. 4, s_0, s, s_1, S_0, and β can all be expressed in terms of $\boldsymbol{\sigma}$ and $\mathbf{v}$. In terms of $\boldsymbol{\sigma}$ the flux terms (4.7) appear very much more complicated, but simplify considerably if one retains only first order relativistic corrections. Equations (4.5) for energy and momentum coupled to the equation of motion (6.38) of Ref. 6 for $\boldsymbol{\sigma}$ are appropriate equations to study if one is interested in the dynamical role of the Thomas precession. Instead, however, we here obtain equations of motion for $\mathbf{s}_1$ and S_2, because they have the general form of conservation laws. To keep the general features of the equations apparent, we do not express s^{μ} and β in terms of s_1 and S_2, as would be necessary if we were looking for solutions. Therefore, we regard (4.7) as a satisfactory expression of the flux in terms of local observables. So substituting (4.7) into (4.5), we get

$$\begin{aligned}
\rho_0 d_t\epsilon &= \rho_0 e\,\mathbf{E}\cdot\mathbf{v} + \partial_t\big\{\rho\, s_0\partial_t\beta - c^{-1}\rho_0\mathbf{v}\cdot\partial_t\mathbf{s}_1\big\}\\
&\qquad -\partial_k\big\{-c\rho\, s^k\partial_t\beta - c\rho_0\partial_t\mathbf{s}\cdot\boldsymbol{\sigma}_k + \rho_0\,(\mathbf{v}\wedge\boldsymbol{\sigma}_k)\cdot\partial_t S_2\big\}\,,
\end{aligned} \tag{4.8a}$$

$$\begin{aligned}
\rho_0 d_t\mathbf{p} &= \rho_0 e\,(\mathbf{E} + c^{-1}\mathbf{v}\times\mathbf{B}) + \partial_t\big\{-\rho\, s_0\nabla\beta + c^{-1}\boldsymbol{\sigma}_j\mathbf{v}\cdot\partial_j\mathbf{s}_1\big\}\\
&\qquad +\partial_k\big\{c\rho\, s^k\nabla\beta + c\rho_0\nabla\mathbf{s}_1\cdot\boldsymbol{\sigma}_k + \rho_0\,\boldsymbol{\sigma}_j(\mathbf{v}\wedge\boldsymbol{\sigma}_k)\cdot\partial_j S_2\big\}\,,
\end{aligned} \tag{4.8b}$$

We now get equations of motion for s_1 and S_2 by taking the relative vector and bivector parts of (4.2a) to get

$$\rho d_\tau\mathbf{s}_1 = -\partial_\mu[M^{\mu}]_1 + \rho[v\wedge p]_1 + [\gamma_\mu\wedge N^{\mu}]_1\,, \tag{4.9a}$$

$$\rho d_\tau S_2 = -\partial_\mu[M^{\mu}]_2 + \rho[v\wedge p]_2 + [\gamma_\mu\wedge N^{\mu}]_2\,. \tag{4.9b}$$

Now to express the right side of (4.9) in relative observables. In the process we keep in mind the facts that $\boldsymbol{\sigma}_k\cdot\mathbf{s} = s^k$ and $\boldsymbol{\sigma}_k\cdot\nabla = \partial_k$ follows from the definitions of $\mathbf{s}$ and ∇ in (6.11) and (6.4) of Ref. 5:

$$\begin{aligned}
c\gamma_\mu\wedge N^{\mu} &= c[\gamma_\mu N^{\mu}]_2 = [\gamma_\mu\gamma_0(c^{-1}N_0^{\mu} - \mathbf{N}^{\mu})]_2\\
&= -\mathbf{N}^0 + c^{-1}\boldsymbol{\sigma}_k N_0^k - \boldsymbol{\sigma}_k\wedge\mathbf{N}^k\,.
\end{aligned}$$

So from (4.7),

$$\begin{aligned}
[\gamma_\mu\wedge N^{\mu}]_1 &= -c^{-1}\mathbf{N}^0 + c^{-2}\boldsymbol{\sigma}_k N_0^k\\
&= -\rho\, s_0\nabla\beta + c^{-1}\rho_0\boldsymbol{\sigma}_j\mathbf{v}\cdot\partial_j\mathbf{s}_1\\
&\qquad - c^{-1}\rho\mathbf{s}\partial_t\beta + c^{-1}\rho_0\{\partial_t\mathbf{s}_1 - c^{-1}\mathbf{v}\cdot\partial_t S_2)\}\,,
\end{aligned} \tag{4.10a}$$

$$
\begin{aligned}
[\gamma_\mu \wedge N^\mu]_2 &= -c^{-1}\boldsymbol{\sigma}_k \wedge \mathbf{N}^k = -c^{-1}\boldsymbol{\sigma}_k \wedge \mathbf{N}_k \\
&= -\rho\mathbf{s}\wedge\nabla\beta - \rho_0\,\nabla\wedge\mathbf{s}_1 + c^{-1}\rho_0\,\boldsymbol{\sigma}_j \wedge (\mathbf{v}\cdot\partial_j S_2)\,.
\end{aligned}
\tag{4.10b}
$$

Also

$$
\begin{aligned}
v\wedge p &= [v_0(1+\mathbf{v}/c)(\epsilon/c - \mathbf{p})]_{(1)+(2)} \\
&= v_0\Big(\frac{\epsilon}{c^2}\,\mathbf{v} - \mathbf{p}\Big) - \frac{v_0}{c}\,\mathbf{v}\wedge\mathbf{p}\,.
\end{aligned}
\tag{4.11}
$$

From (4.2b) we get

$$
\begin{aligned}
M^0 &= \rho\tfrac{1}{2}[\mathbf{s}_1 + S_2, \gamma^0\wedge v] = \rho v_0\{\tfrac{1}{2}[\mathbf{v},\mathbf{s}_1] + \tfrac{1}{2}[\mathbf{v}, S_2]\} \\
&= \rho_0\,\{\mathbf{v}\wedge\mathbf{s}_1 = \mathbf{v}\cdot S_2\}\,,
\end{aligned}
\tag{4.12a}
$$

and, since

$$
\gamma^k\wedge v = [\gamma_0\gamma_0\gamma^k v]_2 = -[\boldsymbol{\sigma}_k v_0(1-\mathbf{v}/c)]_2 = -v_0\{\boldsymbol{\sigma}_k - c^{-1}\boldsymbol{\sigma}_k\wedge\mathbf{v}\}\,,
$$

we also get from (4.2b),

$$
\begin{aligned}
M^k &= -\rho\tfrac{1}{2}[\mathbf{s}_1 + S_2, \boldsymbol{\sigma}_k - c^{-1}i\boldsymbol{\sigma}_k\times\mathbf{v}] \\
&= \rho_0\,\{\,\boldsymbol{\sigma}_k\wedge\mathbf{s}_1 + \boldsymbol{\sigma}_k\cdot S_2 - c^{-1}\mathbf{s}_1\times(\boldsymbol{\sigma}_k\times\mathbf{v}) - c^{-1}\tfrac{1}{2}[\boldsymbol{\sigma}_k\wedge\mathbf{v}, S_2]\}\,.
\end{aligned}
\tag{4.12b}
$$

Hence,

$$
\begin{aligned}
c\partial_\mu[M^\mu]_1 &= \partial_t(\rho_0\mathbf{v}\cdot S_2) + c\nabla\cdot(\rho_0 S_2) - \partial_k\big(\rho_0\mathbf{s}_1\times(\boldsymbol{\sigma}_k\times\mathbf{v})\big)\,, \\
c\partial_\mu[M^\mu]_2 &= \partial_t(\rho_0\mathbf{v}\wedge S_2) + c\nabla\wedge(\rho_0\mathbf{s}_1) - \partial_k\big(\rho_0\tfrac{1}{2}[\boldsymbol{\sigma}_k\wedge\mathbf{v}, S_2]\big)\,.
\end{aligned}
$$

So, at last Eqs. (4.9a,b) can be written

$$
\begin{aligned}
\rho_0 d_t\mathbf{s}_1 &= -\partial_t(\rho\mathbf{v}\cdot S_2) - c\nabla\cdot(\rho_0 S_2) + \partial_k\big(\rho_0\mathbf{s}_1\times(\boldsymbol{\sigma}_k\times\mathbf{v})\big) \\
&\quad + \rho_0(c^{-1}\epsilon\mathbf{v} - c\,\mathbf{p}) - c\rho s_0\nabla\beta + \rho_0\boldsymbol{\sigma}_j\mathbf{v}\cdot\partial_j\mathbf{s}_1 \\
&\quad - \rho\mathbf{s}\partial_t\beta + \rho_0\partial_t\mathbf{s}_1 - c^{-1}\,\rho_0\mathbf{v}\cdot(\partial_t S_2)\,,
\end{aligned}
\tag{4.13a}
$$

$$
\begin{aligned}
\rho_0 d_t S_2 &= -\partial_t(\rho_0\mathbf{v}\wedge\mathbf{s}_1) - c\nabla\wedge(\rho_0\mathbf{s}_1) + \partial_k\big(\rho_0\tfrac{1}{2}[\boldsymbol{\sigma}_k\wedge\mathbf{v}, S_2]\big) \\
&\quad - \rho_0\mathbf{v}\wedge\mathbf{p} - c\rho\mathbf{s}\wedge\nabla\beta - c\rho_0\nabla\wedge\mathbf{s}_1 \\
&\quad + \rho_0\boldsymbol{\sigma}_j\wedge(\mathbf{v}\cdot\partial_j S_2)\,.
\end{aligned}
\tag{4.13b}
$$

With Eqs. (4.8a,b) and (4.13a,b) we have completed the formulation of the Dirac hydrodynamic equations in terms of relative observables. No one can fail to notice how much more complicated these equations are than their proper counterparts (4.1a,b) and (4.2a,b). So, for most purposes it is clearly best to deal with the proper equations.

The relative hydrodynamic equations become much simpler in the N. R. limit. To attain this limit, as we saw in Sec. 7 of Ref. 5, we need only express all spins in terms of S or $\mathbf{s}$ with the identifications

$$
S_2 = S = i\mathbf{s}\,, \quad c\mathbf{s}_1 = \mathbf{v}\cdot S, \quad c s_0 = \mathbf{s}\cdot\mathbf{v}\,,
$$

take $\rho_0 = \rho$, and regarding $c|\,\boldsymbol{\beta}\,| \approx |\,\mathbf{v}\,|$, neglect all terms of relative order c^{-1} or less. Then from (4.8b) we easily get the momentum conservation equation

$$
\rho\, d_t\mathbf{p} = \rho\, e\,(\mathbf{E} + c^{-1}\mathbf{v}\times\mathbf{B}) - \partial_k\mathbf{N}_k\,,
\tag{4.14a}
$$

where the momentum flux $\mathbf{N}_k$ is the limit of (4.7d)

$$
\mathbf{N}_k = \rho c s^k\nabla\beta + \boldsymbol{\sigma}_j\rho S\cdot(\partial_j\mathbf{v}\wedge\boldsymbol{\sigma}_k)\,,
\tag{4.14b}
$$

and from (4.9b) we get the angular momentum conservation equation in terms of the spin

$$\rho d_t S = \rho \mathbf{p} \wedge \mathbf{v}\boldsymbol{\sigma}_k \wedge \mathbf{N}_k - \partial_k M_k \,, \tag{4.15a}$$

where the spin flux M_k is given by

$$M_k \equiv \rho \tfrac{1}{2}[S, \boldsymbol{\sigma}_k \wedge \mathbf{v}] - \rho \boldsymbol{\sigma}_k \wedge (S \cdot \mathbf{v}) \,. \tag{4.15b}$$

In addition, from (4.8a) we get the energy conservation equation

$$\rho d_t S = e\rho \mathbf{E} \cdot \mathbf{v} + \partial_k \big\{ c\rho s_k \partial_t \beta - S \cdot (\partial_t \mathbf{v} \wedge \boldsymbol{\sigma}_k) \big\} \,. \tag{4.16}$$

As already mentioned in the last section, these are the hydrodynamic equations of the Pauli theory.

As an application of the relative hydrodynamic equations we derive a virial theorem for the Dirac theory, though actually it is hardly more difficult to derive the theorem directly from the proper hydrodynamic equations. Differentiating $\mathbf{x} \cdot \mathbf{p}$ and using (4.5b), we get

$$d_t(\mathbf{x} \cdot \mathbf{p}) = \mathbf{v} \cdot \mathbf{p} + \mathbf{x} \cdot [\, e(\mathbf{E} + c^{-1}\mathbf{v} \times \mathbf{B}) - \rho_0^{-1} \partial_\mu \mathbf{N}^\mu \,] \,.$$

But, by (4.7d)

$$\begin{aligned} -\mathbf{x} \cdot (\partial_\mu \mathbf{N}^\mu) &= -\partial_\mu (\mathbf{x} \cdot \mathbf{N}^\mu) + \boldsymbol{\sigma}_k \cdot \mathbf{N}^k \\ &= -\partial_\mu (\mathbf{x} \cdot \mathbf{N}^\mu) + \boldsymbol{\sigma}_k \cdot \mathbf{N}^k + c\rho \mathbf{s} \cdot \nabla \beta - \rho v_0 \big(c \nabla \cdot \mathbf{s}_1 + \mathbf{v} \cdot (\nabla \cdot S_2) \big) \,. \end{aligned}$$

So

$$\begin{aligned} \rho_0 d_t (\mathbf{x} \cdot \mathbf{p}) = \rho_0 \mathbf{v} \cdot \mathbf{p} + e\rho_0 \mathbf{x} \cdot (\mathbf{E} + c^{-1}\mathbf{v} \times \mathbf{B}) \partial_\mu (\mathbf{x} \cdot \mathbf{N}^\mu) \\ \rho_0 \Big(\frac{c}{v_0} \mathbf{s} \cdot \nabla \beta - c \nabla \cdot \mathbf{s}_1 - \mathbf{v} \cdot (\nabla \cdot S_2) \Big) \,, \end{aligned}$$

and since

$$\int d^3x \rho_0 d_t (\mathbf{x} \cdot \mathbf{p}) = \partial_t \int d^3x \rho_0 \, \mathbf{x} \cdot \mathbf{p} = \partial_t \langle\, \mathbf{x} \cdot \mathbf{p} \,\rangle \,,$$

we have

$$\begin{aligned} \partial_t \langle\, \mathbf{x} \cdot \mathbf{p} \,\rangle = \langle\, \mathbf{v} \cdot \mathbf{p} \,\rangle + e \langle\, \mathbf{x} (\mathbf{E} + c^{-1}\mathbf{v} \times \mathbf{B}) \,\rangle - \partial_t \Big\langle \Big(\frac{\mathbf{x} \cdot \mathbf{N}^0}{c\rho_0} \Big) \Big\rangle \\ + \Big\langle \Big(\frac{c}{v_0} \mathbf{s} \cdot \nabla \beta - c \nabla \cdot \mathbf{s}_1 - \mathbf{v} \cdot S_2 \Big) \Big\rangle \,. \end{aligned} \tag{4.17}$$

Hence, for stationary states we have

$$\langle\, \mathbf{v} \cdot \mathbf{p} \,\rangle + e \langle\, \mathbf{x} (\mathbf{E} + c^{-1}\mathbf{v} \times \mathbf{B}) \,\rangle = - \Big\langle \Big(\frac{c}{v_0} \mathbf{s} \cdot \nabla \beta - c \nabla \cdot \mathbf{s}_1 - \mathbf{v} \cdot (\nabla \cdot S_2) \Big) \Big\rangle \,. \tag{4.18}$$

This can be related to the energy by taking the expectation value of (6.30) in Ref. 5 to get

$$\langle\, E \,\rangle = \Big\langle\, c \frac{\overline{\Omega} \cdot S}{v_0} \,\Big\rangle + mc^2 \Big\langle\, \frac{\cos \beta}{v_0} \,\Big\rangle + \langle\, \mathbf{v} \cdot \mathbf{p} \,\rangle + \langle\, V \,\rangle \,. \tag{4.19}$$

Now we recall from (6.27a) in Ref. 5 that $\overline{\Omega} = -\Box \wedge \mathbf{v} + v \cdot (i \Box \beta)$. To express $\overline{\Omega} \cdot S$ in terms of relative variables, note that the divergence of $v \cdot S = 0$ gives

$$(\Box \wedge v) \cdot S = (v \wedge \Box) \cdot S = (v \Box S)_{(0)} \,.$$

Hence

$$-(\Box\wedge v)\cdot S = -v_0[(1+\mathbf{v}/c)(\partial_0+\nabla)(\mathbf{s}_1+S_2)]_{(0)}$$
$$= -\frac{v_0}{c}\left[\mathbf{v}\cdot\partial_0\mathbf{s}_1 + c\nabla\cdot\mathbf{s}_1 + \mathbf{v}\cdot(\nabla\cdot S_2)\right].$$

Also,

$$S\cdot\big(v\cdot(i\Box\beta)\big) = [Svi\Box\beta]_{(0)} = s\cdot\Box\beta = s_0\partial_0 + \mathbf{s}\cdot\nabla\beta\,.$$

Hence, we have the general formula,

$$c\rho\overline{\Omega}\cdot S = \rho_0\Big\{\frac{cs_0}{v_0}\,\partial_0\beta - \mathbf{v}\cdot\partial_0\mathbf{s}_1 + \frac{c}{v_0}\,\mathbf{s}\cdot\nabla\beta - c\nabla\mathbf{s}_1 - \mathbf{v}\cdot(\nabla\cdot S_2)\Big\}\,. \tag{4.20}$$

Taking the expectation value of (4.20) and comparing with (4.18), we find *for stationary states* Hence, for stationary states we have

$$\left\langle \frac{c\overline{\Omega}\cdot S}{v_0}\right\rangle = \left\langle \Big(\frac{c}{v_0}\,\mathbf{s}\cdot\nabla\beta - c\nabla\cdot\mathbf{s}_1 - \mathbf{v}\cdot(\nabla\cdot S_2)\Big)\right\rangle$$
$$= -\langle\,\mathbf{v}\cdot\mathbf{p}\,\rangle - e\langle\,\mathbf{x}\cdot(\mathbf{E}+c^{-1}\mathbf{v}\times\mathbf{B})\,\rangle = \left\langle \frac{c\varpi^L\cdot\Sigma}{v_0}\right\rangle\,. \tag{4.21}$$

Recalling that the last term in (4.21) is equal to the first by (6.17) and (6.49) of Ref. 5; the factor c/v_0 would be missing, as (6.38) of Ref. 4 shows, if ϖ^L were defined as the angular velocity corresponding to the total time derivative instead of the proper time derivative. Thus, (4.21) is exactly a virial theorem for the generalized Larmor precession energy, or as it was called in Sec. 6 of Ref. 4, the *internal energy.*

Substituting (4.21) in (4.19), we obtain the result that for any stationary state in the Dirac theory

$$\langle\,E\,\rangle = mc^2\left\langle\frac{\cos\beta}{v_0}\right\rangle + \langle\,V\,\rangle - e\,\langle\,\mathbf{x}\cdot(\mathbf{E}+c^{-1}\mathbf{v}\times\mathbf{B})\,\rangle\,. \tag{4.22}$$

A Coulomb field has the special property that

$$e\mathbf{x}\cdot\mathbf{E} = -\mathbf{x}\cdot\nabla V = V\,. \tag{4.23}$$

Hence for a Coulomb field alone, (4.22) reduces to the simple form

$$\langle\,E\,\rangle = mc^2\left\langle\frac{\cos\beta}{v_0}\right\rangle + \langle\,V\,\rangle = mc^2\!\int d^3x\,\rho\cos\beta\,. \tag{4.24}$$

This ought to tell us something important about the interpretation of β but we do not know what. At least we can use it as additional support for our contention that β must, on the average, be a small quantity. Thus, if $\langle\,E\,\rangle$ does not deviate much from mc^2, then, since $v_0 \geq 1$ everywhere, (4.29) implies that $\langle\cos\beta\,\rangle \approx 1$. This being true, tI-Uf, we can expand v_0 and $\cos\beta$ in (4.24) to get the approximate expression

$$\langle\,E\,\rangle - mc^2 \approx -\langle\,\tfrac{1}{2}mv^2\,\rangle - \tfrac{1}{2}mc^2\langle\,\beta^2\,\rangle\,. \tag{4.25}$$

This, as should be expected, is just the virial theorem one obtains by using the Pauli theory. It should be compared with

$$\langle\,E\,\rangle - mc^2 \approx \langle\,\tfrac{1}{2}m\boldsymbol{\sigma}^2\,\rangle + \langle\,\tfrac{1}{2}mc^2\beta^2\,\rangle + \langle\,V\,\rangle\,. \tag{4.26}$$

which is obtained immediately by integrating (7.5) of Ref. 5.

*Some of this work appeared in a thesis, Arizona State University, 1972.

[1]D. Blokhgintsev, *Principles of Quantum Mechanics* (Allyn and Bacon, Boston, 1964), Sec. 61.
[2]G. Breit, *Phys. Rev.* **53**, 153 (1938).
[3]W. Gordon, *Z. Phys.* **50**, 360 (1928).

[4]D. Hestenes, *J. Math. Phys.* **14**, 893 (1973). [By an error in proof, the physical interpretations of the Dirac and Gordon currents in (5.6) of this reference got interchanged in the subsequent sentence.]

[5]D. Hestenes, *J. Math. Phys.* **16**, 556 (1975). Preceding paper.

[6]D. Hestenes and R. Gurtler, *Am. J. Phys.* **39**, 1028 (1971).

[7]E. Madelung, *Z. Phys.* **40**, 322 (1926).

[8]W. Wilhelm, *Phys. Rev.* D **1**, 2278 (1970).

[9]T. Takabayasi, *Prog. Theor. Phys.* **14**, 228 (1955); *Suppl. Prog. Theor. Phys.* (**4**), 1 (1957).

[10]O. Costa de Beauregard, *J. Math. Pures Appl.* **22**, 85 (1943).

[11]D. Hestenes, *J. Math. Phys.* **15**, 1768 (1974).

[12]E. Hill and R. Landshoff, *Rev. Mod. Phys.* **10**, 87 (1938).

[13]H. Bethe and E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer-Verlag, Berlin, 1957), p.54.

[14]R. Feynman, *Quantum Electrodynamics* (Benjamin, New York, 1961), p. 51.

[15]L. Foldy and S. Woutheysen, *Phys. Rev.* **78**, 29 (1950).

[16]W. Barker and F. Glover, *Phys. Rev.* **99**, 317 (1955).

[17]D. Hestenes, *Space-Time Algebra* (Gordon and Breach, New York, 1966).