

The Pauli and Lévy-Leblond equations, and the spin current density

James M Wilkes 

2127 NW Tilia Trail, Stuart, FL 34994, United States of America

E-mail: mike.wilkes06@gmail.com

Received 30 October 2019, revised 26 January 2020

Accepted for publication 10 February 2020

Published 27 March 2020



CrossMark

Abstract

We review the literature on the Pauli equation and its current density, discussing the progression from the original phenomenological version of Pauli to its derivation by Lévy-Leblond from a linearization of the Schrödinger equation. It was established conclusively by Lévy-Leblond's work that the spin of a spin-1/2 particle such as an electron is non-relativistic in nature, contrary to what was often stated following Dirac's derivation of a relativistic wave equation, and his subsequent demonstration that Pauli's spin interaction term appeared in the non-relativistic limit. In this limit, the Gordon decomposition of the associated probability current density was found to contain a spin-dependent term. Such a term does not follow, however, from the usual derivation of the current density from the Pauli equation, although various physically motivated but otherwise ad hoc explanations were put forward to account for it. We comment on the only exception to these of which we are aware implying the spin term in the current was in fact non-relativistic in nature. However, the earlier work of Lévy-Leblond had already shown, with no additional assumptions, that this term was a prominent feature of the current density derived from his equation. Hence, just as with the spin itself, the spin current was non-relativistic, claims to the contrary notwithstanding. We present a somewhat simplified derivation of the Lévy-Leblond equation and its current density, commenting on possibilities for experimental work that might indicate measurable consequences of the spin term in the current density. The presentation is at a level of quantum mechanics, electromagnetic theory, and mathematical methods typical of courses taken by first-year graduate students in physics, assuming some familiarity with Pauli spin matrices and the Dirac equation. Calculations are presented in detail, making the material suitable for use by teachers as supplementary problems for assignments, and development of theoretical concepts peculiar to the Pauli equation and its probability current density.

Keywords: Pauli equation, Lévy-Leblond equation, spin current density, spin-1/2 particles

1. Introduction

Pauli's equation for the electron is a topic of general interest in quantum mechanics textbooks [1–3] and journal articles [4, 5]. Although it explains many of the experimental results associated with non-relativistic electrons, its probability current density turns out to include an unexpected additional spin-dependent term (the spin current), which has received considerable attention [5–9]. We begin with a selective review of some of the relevant earlier work.

As background, in 1927 Pauli [10] introduced his famous spin matrices in modifying the non-relativistic Schrödinger equation to account for Goudsmit and Uhlenbeck's [11, 12] hypothesis that spectral data explainable by half-integral quantum numbers implied a *half-integer spin angular momentum* for the electron. His ansatz was to add a phenomenological term to the usual Hamiltonian for an electron moving in an electromagnetic field, viz., the interaction energy of a magnetic field with an electron magnetic moment proportional to its intrinsic spin angular momentum. Introduction of spin matrices to describe this spin angular momentum necessitated replacing the complex scalar wave function of the Schrödinger equation by a two-component spinor wave function. The standard form of Pauli's equation as given in, for example, [2] (converted to SI units, used throughout the present work), is

$$\begin{aligned} & \left[\frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 \mathbf{1} + q\phi \mathbf{1} - \hat{\boldsymbol{\mu}}_s \cdot \mathbf{B} \right] \psi(\mathbf{r}, t) \\ & = \hat{E} \mathbf{1} \psi(\mathbf{r}, t), \end{aligned} \quad (1)$$

where $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ is a 2-component spinor, $\mathbf{1}$ is the 2×2 identity matrix (often omitted, but to be understood), m and $q (= -e)$ are the mass and charge of the electron, and $\hat{\mathbf{p}} = -i\hbar \nabla$ and $\hat{E} = i\hbar \frac{\partial}{\partial t}$ are the usual linear momentum and energy operators. The appearance of (real-valued) electromagnetic scalar and vector potentials, $\phi(\mathbf{r}, t)$ and $\mathbf{A}(\mathbf{r}, t)$, is a consequence of using the gauge invariant *minimal coupling* assumption (see, for example, [2], p 205) to describe the interaction with external magnetic and electric fields defined by $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\nabla\phi - \frac{\partial}{\partial t}\mathbf{A}$. Finally,

$$\hat{\boldsymbol{\mu}}_s = \left(\frac{q\hbar}{2m} \right) \boldsymbol{\sigma} \quad (2)$$

is the electron magnetic moment operator, where $\boldsymbol{\sigma}$ is the Pauli spin vector with Pauli spin matrices σ_i , $i = 1, 2, 3$, as components.

The following year Dirac [13] presented his relativistic wave equation for a free electron, and by making use of the minimal coupling replacement to include electromagnetic interactions showed that it contained two terms that 'can be regarded as the additional potential energy of the electron due to its new degree of freedom'. One of these was, of course, the term involving the electron magnetic moment interacting with a magnetic field, introduced 'by hand' in Pauli's equation. This prediction, requiring no ad hoc assumptions, was a triumph of the Dirac theory.

After Dirac's discovery it became common lore in many physics textbooks (see, for example, the extensive list of quotations compiled by Lévy-Leblond [14]) to regard electron spin as a relativistic phenomenon, although there were prominent exceptions to this stance ([14], p 141). Among these exceptions, e.g. [4, 15–17], the observation was made that the spin 1/2 property could be incorporated into the spin-0 non-relativistic Schrödinger equation by simply replacing the kinetic momentum operator $\hat{\pi} = \hat{p} - q\mathbf{A}$ appearing there with the product $\boldsymbol{\sigma} \cdot \hat{\pi}$. To see how this is possible, we follow Sakurai [16] by beginning with the spin-0 non-relativistic *free-electron* Schrödinger equation of wave mechanics:

$$\frac{\hat{p}^2}{2m}\Psi(\mathbf{r}, t) = \hat{E}\Psi(\mathbf{r}, t), \quad (3)$$

where Ψ is Schrödinger's complex scalar wave function. Recall the important spin-vector identity

$$(\boldsymbol{\sigma} \cdot \hat{\mathbf{a}})(\boldsymbol{\sigma} \cdot \hat{\mathbf{b}}) = (\hat{\mathbf{a}} \cdot \hat{\mathbf{b}})\mathbf{1} + i\boldsymbol{\sigma} \cdot (\hat{\mathbf{a}} \times \hat{\mathbf{b}}), \quad (4)$$

where $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are any two vector operators that commute with $\boldsymbol{\sigma}$. For $\hat{\mathbf{a}} = \hat{\mathbf{b}} = \hat{p}$ one finds from equation (4) the operator identity $(\boldsymbol{\sigma} \cdot \hat{p})^2 = \hat{p}^2\mathbf{1}$. Sakurai refers to the theory that results from making this substitution in Schrödinger's equation (3), and replacing the wave function Ψ by a 2-component spinor ψ , as the Schrödinger–Pauli theory ([16], footnote, p 79), i.e.,

$$\frac{(\boldsymbol{\sigma} \cdot \hat{p})^2}{2m}\psi(\mathbf{r}, t) = \hat{E}\psi(\mathbf{r}, t), \quad (5)$$

which we refer to as the *Schrödinger–Pauli* equation. It has the plane wave solutions $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\chi$, where χ is a constant but otherwise arbitrary *spinor*, giving rise to the dispersion relation $\omega = \hbar\mathbf{k}^2/2m$ characteristic of non-relativistic particles. The effects of spin manifest by allowing the electron to interact with an electromagnetic field. This is accomplished by introduction of the gauge invariant minimal coupling replacements of \hat{p} with $\hat{\pi} = \hat{p} - q\mathbf{A}$, and \hat{E} with $\hat{\varepsilon} = \hat{E} - q\phi$, in which case the Schrödinger–Pauli equation (5) takes the form

$$\frac{(\boldsymbol{\sigma} \cdot \hat{\pi})^2}{2m}\psi(\mathbf{r}, t) = \hat{\varepsilon}\mathbf{1}\psi(\mathbf{r}, t). \quad (6)$$

Applying the identity (4) with $\hat{\mathbf{a}} = \hat{\mathbf{b}} = \hat{\pi} = \hat{p} - q\mathbf{A}$, we have in the numerator of the last line:

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \hat{\pi})^2\psi &= \hat{\pi}^2\mathbf{1}\psi + i\boldsymbol{\sigma} \cdot [(\hat{p} - q\mathbf{A}) \times (\hat{p} - q\mathbf{A})]\psi, \\ &= \hat{\pi}^2\mathbf{1}\psi - q\hbar\boldsymbol{\sigma} \cdot [\nabla \times (\mathbf{A}\psi) + \mathbf{A} \times (\nabla\psi)], \\ &= \hat{\pi}^2\mathbf{1}\psi - q\hbar(\boldsymbol{\sigma} \cdot \mathbf{B})\psi, \end{aligned} \quad (7)$$

where the cross products $\hat{p} \times \hat{p}$ and $\mathbf{A} \times \mathbf{A}$ both give vanishing contributions, we have replaced $\hat{p} = -i\hbar\nabla$, and used the well-known vector identity $\nabla \times (\mathbf{A}\psi) = \nabla\psi \times \mathbf{A} + (\nabla \times \mathbf{A})\psi$, where $\nabla \times \mathbf{A} = \mathbf{B}$. Substituting equation (7) together with $\hat{\varepsilon} = \hat{E} - q\phi$ in equation (6) shows that the *minimally coupled* Schrödinger–Pauli equation (6) reproduces precisely the Pauli equation (1). The result follows without regard for the Dirac equation, obtaining Pauli's spin interaction term directly from the minimally coupled Schrödinger equation and the replacement of $\hat{\pi}$ by $\boldsymbol{\sigma} \cdot \hat{\pi}$, as suggested by Feynman [15].

The conclusion, then, is that the Pauli spin interaction term is an intrinsically non-relativistic effect. Still, one may be left with the feeling that if not for the purely mathematical and fortuitous existence of the spin identity, the Pauli equation would seem a slightly strange concoction forced upon the Schrödinger equation. This situation is somewhat alleviated by

the fact that the minimally coupled Schrödinger–Pauli equation, from which Pauli’s equation follows, can be shown [3, 16, 20] to be the *non-relativistic limit* of Dirac’s equation. Such a derivation is convincing, yet fails to satisfy the desire for a purely non-relativistic account. The work of Lévy-Leblond, to be discussed later, seeks to fill this gap.

2. The missing term in the Pauli current density

As with the Schrödinger equation, one would hope to determine from the Pauli equation a probability density ρ and current \mathbf{J}' satisfying the probability conservation (or continuity) equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J}' = 0. \quad (8)$$

Such derivations are offered, for example, in [1, 2], and [4–7]. A canonical derivation of the current density from Pauli’s equation is that of [2], p 340. Defining the probability density by $\rho = \psi^\dagger \psi$, where $\psi^\dagger = (\psi_1^* \psi_2^*)$ denotes the Hermitian conjugate or adjoint of ψ , the current density is found there to be given (in our notation and SI units) by

$$\mathbf{J}' = -\frac{i\hbar}{2m} [\psi^\dagger (\nabla \psi) - (\nabla \psi)^\dagger \psi] - \frac{q}{m} \mathbf{A} \psi^\dagger \psi. \quad (9)$$

Since the Pauli equation is the non-relativistic limit of Dirac’s equation, one might expect that equation (9) would be the corresponding non-relativistic limit of the current density of Dirac’s equation, but such is not the case. Relativistic quantum mechanics shows that in this limit there appears an additional term, referred to as the spin current density. The details are given, for example, in Sakurai’s discussion [16], p 107, of the Gordon decomposition of the Dirac current and its non-relativistic limit. It is found (see also [6]) that this limit yields instead the following expression:

$$\mathbf{J} = \mathbf{J}' + \mathbf{J}_{\text{spin}}, \quad (10)$$

where

$$\mathbf{J}_{\text{spin}} \equiv \frac{\hbar}{2m} \nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi) \quad (11)$$

is the additional *spin current density*. This term must be a feature of the current density for *any* spin-1/2 particle described by the Dirac equation, even when the particle carries no charge, e.g. a neutron described by the Dirac equation with the addition of a non-minimal coupling term (as in [20], p 241). It is important to note, with Nowakowski [6], that this spin current is also manifest in the reduction of the Dirac equation for a *free particle*, i.e. in the *absence* of any interactions.

Various writers have suggested methods for supplying the spin current missing from equation (9). For example, in the venerable and oft-cited textbook of Landau and Lifshitz [21], the problem of a charged particle moving in a magnetic field is considered. A variational principle is applied, varying the expectation value of the Pauli Hamiltonian of equation (1) (ignoring the scalar potential) under a variation in the vector potential \mathbf{A} . From this they derive the *charge* current density $q\mathbf{J}$, where \mathbf{J} is given precisely by equation (10), including the spin current. Without the minimal coupling to a magnetic field via \mathbf{A} and the charge q , however, such a derivation apparently would fail to provide the desired spin current term, which we know from equation (10) should appear even in the absence of such an interaction.

Of other works known to the present author, each relies in some way upon the following line of reasoning. Since the continuity equation (8) involves only the divergence of the current density, it is always possible to add a divergence-free term to the current density without changing the equation. The spin current proportional to a curl is, of course, just such a term. But since the Pauli equation does not provide this, one must rely on reasonable *physical* arguments to supply it, which we note should hold *even in the absence of interaction potentials*.

As an example, Mita [7] takes the novel approach of beginning with the expectation value of the spin angular momentum $\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}$, and showing that it can be written as the volume integral of an angular momentum, viz.,

$$\langle \mathbf{S} \rangle = m \int_{V_0} \mathbf{r} \times \mathbf{j}_s \, d^3r, \quad \mathbf{j}_s = \frac{\hbar}{4m} \nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi). \quad (12)$$

Mita considers this \mathbf{j}_s to be ‘the virtual probability current density that gives rise to the spin angular momentum of the electron.’ He also notes, however, that ‘the proper gyromagnetic ratio of the particle cannot be obtained in the context of our analysis based on nonrelativistic quantum mechanics.’ In fact, his current density gives a gyromagnetic ratio that is half the correct value.

Greiner [2] and Parker [5], as well as Hodge *et al* [9] (following Greiner), call upon a result from electromagnetic theory, viz., that a magnetization field \mathbf{M} induces an additional (electric) current density $q\mathbf{J}_M = \nabla \times \mathbf{M}$ (see, for example [18, 19]). The expectation value of the electron magnetic moment is

$$\langle \hat{\boldsymbol{\mu}}_s \rangle = \int \psi^\dagger \hat{\boldsymbol{\mu}}_s \psi \, d^3r = \frac{q\hbar}{2m} \int \psi^\dagger \boldsymbol{\sigma} \psi \, d^3r, \quad (13)$$

the integrand of equation (13) defining the magnetic moment density, which these authors *identify* with the magnetization \mathbf{M} of electromagnetic theory. One thus finds for the associated probability current density:

$$\mathbf{J}_M = \frac{1}{q} \nabla \times \mathbf{M} = \frac{\hbar}{2m} \nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi), \quad (14)$$

the correct expression for the spin current. As admitted in [9], referring to the current density, ‘This argument is not rigorous, but it does hint that another term should be added’

As a final example, we consider the work of Shikakhwa *et al* [8]. They discuss Nowakowski’s [6] spin current term, denoting it by \mathbf{J}_M and referring to it as a magnetization current. These authors are clearly aware that the spin current cannot be derived from the Pauli equation (1) in the usual way (as, for example, in [2], p 340), so they instead begin with the minimally coupled Schrödinger–Pauli equation (6), ignoring the scalar potential term:

$$\frac{(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2}{2m} \psi = i\hbar \frac{\partial \psi}{\partial t}. \quad (15)$$

The key to the success of their approach is to use equation (15) in deriving a probability conservation equation *before* expanding $(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2$ (see equation (7)), since otherwise the Pauli equation (1) follows. Note that although their goal is to obtain Nowakowski’s expression for the non-relativistic limit of the current density for a *free* spin-1/2 particle, given in their equation (2), they actually work with equation (15), above, for such a particle *interacting* with a magnetic field. However, they emphasize more than once their use of $\hat{\boldsymbol{\pi}}$ as merely a convenient bookkeeping device in arriving at the desired result; the vector potential is ignored in the final step of their derivation. Their motivation for using $\hat{\boldsymbol{\pi}}$ rather than $\hat{\boldsymbol{p}}$, which naturally

appears in the free particle Hamiltonian, is the concern that using $\hat{\mathbf{p}}$ in their derivation would be problematic due to the commutation of its components. We show in the next section that this concern can be addressed by a different derivation that avoids the problem.

In concluding our (in no way comprehensive) review of attempts to derive the spin current term we draw attention to remarks of Shikakhwa *et al* [8] in the paragraph following equation (12c) of their paper, which is their final result for the current density. Referring to the derivation of this equation, they conclude: ‘Nevertheless, the fact that we can find the correct term \mathbf{J}_M by not dropping a term that gives a vanishing contribution in the continuity equation adds credence to the current approach.’ Note that their \mathbf{J}_M is the spin current density that we denote by \mathbf{J}_{spin} .

This last sentence we feel is the relevant conclusion of their paper. They have obtained the current density, including the spin current term, in a form that is precisely the *correct non-relativistic limit* of the Dirac current density.

However, prior to this conclusion we read: ‘It is possible to continue working with equation (10) and to drop the terms that give vanishing contributions in $\nabla \cdot \mathbf{J}_M = 0$. In such an approach we can obtain only the \mathbf{J}_0 contribution to the current. These kinds of ambiguities are expected because we are not following the only correct methodology (that is, using a covariance argument in the Dirac equation).’ Now, we know that since the spin current is proportional to a curl, it will not make a contribution to the *probability conservation equation*; only the familiar convection term \mathbf{J}' (denoted by \mathbf{J}_0 in [8] in the absence of external electromagnetic field interactions) contributes to probability conservation. And, *if the \mathbf{J}_{spin} term is dropped*, certainly all that is left in the current density in this case is \mathbf{J}' . But why would one consider *dropping* the \mathbf{J}_{spin} term after finding it to be present in the derivation of the current density? Nevertheless, it is clear from their discussion that the authors feel ambiguities remain in their expression for the current density.

Their concern calls attention to the question of uniqueness of their result. In this regard, we suggest that the work of Holland [22, 23] and Holland and Philippidis [24] goes far to answer this, and to perhaps alleviate their concern. In particular, Holland [23] concludes under very general conditions that, first, the Dirac current density is the unique Lorentz covariant current of Dirac’s relativistic equation; second, since the Pauli equation is the non-relativistic limit of the Dirac equation, the Pauli current density should be the non-relativistic limit of the Dirac current; and third, ‘the Pauli current so obtained will be unique since we have already proved its relativistic forebear to be unique’. It seems fair to summarize these results by saying that Lorentz covariance assures the uniqueness of the relativistic Dirac current density, while agreement with the non-relativistic *limit* of this current density assures the uniqueness of the non-relativistic current density. Holland’s [23] work suggests that Shikakhwa *et al* [8] have succeeded in deriving *unambiguously* the correct spin 1/2 current density from the non-relativistic minimally coupled Schrödinger–Pauli equation of equation (15), although just as with the spin, relativity is invaluable in confirming its existence and uniqueness. The situation is thus reminiscent of the claim by early proponents of the Dirac equation that relativity was necessary to explain the electron’s spin and its correct spin magnetic moment. The work of Shikakhwa *et al* [8] suggests that the claim that relativity is required to explain the *spin current density* is similarly unfounded.

3. Derivation of the spin current density from the Schrödinger–Pauli equation for a free spin-1/2 particle

The authors of [8] derived the non-relativistic current density for a free spin-1/2 particle beginning with the minimally coupled Schrödinger–Pauli equation, and later set $\mathbf{A} = \mathbf{0}$ in the result. We show here that the current density can be derived directly from the free-particle Schrödinger–Pauli equation (5). From this equation we have

$$\frac{\partial \psi}{\partial t} = -\frac{i}{2m\hbar}(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2 \psi,$$

and its Hermitian conjugate equation:

$$\frac{\partial \psi^\dagger}{\partial t} = \frac{i}{2m\hbar}[(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2 \psi]^\dagger.$$

Taking $\rho = \psi^\dagger \psi$ as the probability density, we find for its time derivative

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \psi^\dagger \frac{\partial \psi}{\partial t} + \frac{\partial \psi^\dagger}{\partial t} \psi, \\ &= -\frac{i}{2m\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2 \psi] - [(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2 \psi]^\dagger \psi \}. \end{aligned} \quad (16)$$

At this point, our derivation departs from the one given in [8], where our momentum operator $\hat{\mathbf{p}}$ in equation (16) was replaced by the kinetic momentum $\hat{\boldsymbol{\pi}}$ in their corresponding equation (8b). We instead retain $\hat{\mathbf{p}}$ and introduce an auxiliary spinor χ defined by

$$\chi = -\frac{1}{2mc}(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})\psi, \quad (17)$$

the speed of light c included only to ensure dimensional consistency (in no way to be construed as introducing a relativistic concept into the theory). Equation (16) can then be written as

$$\begin{aligned} \frac{1}{c} \frac{\partial \rho}{\partial t} &= \frac{i}{\hbar} \psi^\dagger [(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})\chi] - \frac{i}{\hbar} [(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})\chi]^\dagger \psi, \\ &= \psi^\dagger (\boldsymbol{\sigma} \cdot \nabla \chi) + (\boldsymbol{\sigma} \cdot \nabla \chi)^\dagger \psi, \end{aligned}$$

The divergence identity

$$\begin{aligned} \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi) &= \psi^\dagger (\boldsymbol{\sigma} \cdot \nabla \chi) + (\boldsymbol{\sigma} \cdot \nabla \chi)^\dagger \psi \\ &\quad + (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \chi + \chi^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi), \end{aligned}$$

allows us to replace the right-hand side of the last equation, yielding

$$\frac{1}{c} \frac{\partial \rho}{\partial t} = \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi) - (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \chi - \chi^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi).$$

From equation (17) and its adjoint, however, we have

$$\chi = \frac{i\hbar}{2mc} \boldsymbol{\sigma} \cdot \nabla \psi, \quad \chi^\dagger = -\frac{i\hbar}{2mc} (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger,$$

which, when substituted in the previous equation eliminates the last two terms, bringing it to simply

$$\frac{1}{c} \frac{\partial \rho}{\partial t} = \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi). \quad (18)$$

Substituting for χ and χ^\dagger in (18) (thereby *eliminating* the dimensional factor of c) yields

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{i\hbar}{2m} \{ \nabla \cdot [\psi^\dagger \boldsymbol{\sigma} (\boldsymbol{\sigma} \cdot \nabla \psi) - (\nabla \psi \cdot \boldsymbol{\sigma})^\dagger \boldsymbol{\sigma} \psi] \}, \\ &= -\nabla \cdot \mathbf{J}, \end{aligned} \quad (19)$$

allowing identification of the components of the probability current density \mathbf{J} :

$$J_i = \frac{i\hbar}{2m} [(\nabla_j \psi)^\dagger \sigma_j \sigma_i \psi - \psi^\dagger \sigma_i \sigma_j (\nabla_j \psi)],$$

summation over repeated indices being understood. Using the Pauli matrix identity $\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk} \sigma_k$, the current density components can be written as

$$\begin{aligned} J_i &= \frac{i\hbar}{2m} [(\nabla_j \psi)^\dagger (\delta_{ij} - i\epsilon_{ijk} \sigma_k) \psi - \psi^\dagger (\delta_{ij} + i\epsilon_{ijk} \sigma_k) \nabla_j \psi], \\ &= \frac{i\hbar}{2m} [(\nabla_i \psi)^\dagger \psi - \psi^\dagger (\nabla_i \psi) - i\epsilon_{ijk} \nabla_j (\psi^\dagger \sigma_k \psi)], \end{aligned}$$

or in vector form:

$$\mathbf{J} = -\frac{i\hbar}{2m} [\psi^\dagger (\nabla \psi) - (\nabla \psi)^\dagger \psi] + \frac{\hbar}{2m} [\nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi)], \quad (20)$$

the correct current density for an external interaction-free spin-1/2 particle.

4. The Lévy-Leblond equation

Nearly 40 years after Pauli and Dirac's seminal work, in a paper [25] 'devoted to a detailed study of non-relativistic particles and their properties, as described by Galilei invariant wave equations,' J-M Lévy-Leblond, inspired by Dirac's heuristic derivation [13] of his relativistic wave equation for the electron, obtained from the free-particle Schrödinger equation a non-relativistic analog now referred to as the Lévy-Leblond equation. From it the Pauli equation *follows* as an immediate consequence after calling upon the minimal coupling replacement. In addition, the probability current computed from the Lévy-Leblond equation includes the correct spin current contribution.

A derivation of the Lévy-Leblond equation eventually involves the introduction of a 4-component spinor (bispinor), and 4×4 matrices satisfying the Dirac algebra, which may be considered at too high a level to be introduced in an early course on quantum mechanics. As we show, however, these matrices are easily constructed from the Pauli matrices, so if the Pauli theory has been introduced, there is really no great hurdle to be overcome in using them. Even so, if the decision is made to avoid their use, we should mention that having made plausible the Pauli equation by, e.g. the Sakurai argument leading to equation (6), a coupled set of equations *linear* in the momentum operator, i.e. first order in the spatial derivatives, can be obtained by simply introducing, similar to what was done in the last section, an *auxiliary* spinor χ defined by

$$\chi = -\frac{1}{2mc} [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})] \psi, \quad (21)$$

It then follows from equation (6) that

$$-c[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})] \chi + q\phi \psi = \hat{E} \psi. \quad (22)$$

This pair of equations is precisely the system derived by Lévy-Leblond in [25], that is, the Lévy-Leblond equations follow from the Pauli equation. Lévy-Leblond's distinguished contribution, on the other hand, was to provide a derivation of these equations from the non-relativistic Schrödinger equation, from which Pauli's equation then follows. A derivation based on his insights is offered in what follows.

4.1. Linearizing the Schrödinger equation

We begin with the Schrödinger equation (3) for a free particle of mass m , multiplying both sides by $2m$ to write it in the more convenient form

$$(\hat{p}_i \hat{p}_j \delta_{ij} - 2m\hat{E})\Psi(\mathbf{r}, t) = 0. \quad (23)$$

Following Lévy-Leblond [25], we seek a wave equation linear in both the energy and momenta operators:

$$\hat{\theta} \Psi(\mathbf{r}, t) \equiv \left(\frac{\hat{A}}{c} \hat{E} + \hat{B}_i \hat{p}_j \delta_{ij} + mc \hat{C} \right) \Psi(\mathbf{r}, t) = 0, \quad (24)$$

where \hat{A} , \hat{C} , and \hat{B}_i ($i = 1, 2, 3$) are dimensionless linear operators to be determined. They are assumed to depend on neither the spatial coordinates nor time, hence commute with \hat{E} and each \hat{p}_i , but not necessarily with each other. The factors $1/c$ of \hat{A} , and mc of \hat{B}_i , where c is an arbitrary constant speed, are introduced to insure the correct units of linear momentum in each term. It will later be identified with the speed of light for compatibility with the non-relativistic reduction of the Dirac equation. From equation (24) we obtain, by operating on both sides with $\hat{\theta}$,

$$\hat{\theta}^2 \Psi(\mathbf{r}, t) \equiv \left(\frac{\hat{A}}{c} \hat{E} + \hat{B}_i \hat{p}_j \delta_{ij} + mc \hat{C} \right)^2 \Psi(\mathbf{r}, t) = 0. \quad (25)$$

In doing so we have departed from Lévy-Leblond's [25] and other derivations (e.g. [2], chapter 13, and [26, 27]) by omitting primed counterparts of the operators \hat{A} , \hat{B}_i , and \hat{C} . Conditions are now set on these operators by requiring that equations (23) and (25) agree. Squaring the linear operator, we obtain

$$\begin{aligned} \frac{\hat{A}^2}{c^2} \hat{E}^2 + \frac{1}{c} (\hat{A} \hat{B}_i + \hat{B}_i \hat{A}) \hat{E} \hat{p}_i + m (\hat{A} \hat{C} + \hat{C} \hat{A}) \hat{E} \\ + \frac{1}{2} (\hat{B}_i \hat{B}_j + \hat{B}_j \hat{B}_i) \hat{p}_i \hat{p}_j + mc (\hat{B}_i \hat{C} + \hat{C} \hat{B}_i) \hat{p}_i \\ + m^2 c^2 \hat{C}^2 = \hat{p}_i \hat{p}_j \delta_{ij} - 2m\hat{E}, \end{aligned}$$

where the fourth term on the left-hand side was obtained by interchanging summation indices to rewrite it as a symmetrized sum. This polynomial equation in \hat{E} and the \hat{p}_i is satisfied for all \hat{A} , \hat{C} , and \hat{B}_i if and only if each of its coefficients is zero, yielding

$$\begin{aligned}
\hat{A}^2 = \hat{0}, \quad \hat{C}^2 = \hat{0}, \quad \hat{A}\hat{C} + \hat{C}\hat{A} = -2\hat{I}, \\
\hat{A}\hat{B}_i + \hat{B}_i\hat{A} = \hat{0}, \quad \hat{C}\hat{B}_i + \hat{B}_i\hat{C} = \hat{0}, \\
\hat{B}_i\hat{B}_j + \hat{B}_j\hat{B}_i = 2\delta_{ij}\hat{I}, \quad (i, j = 1, 2, 3).
\end{aligned} \tag{26}$$

where \hat{I} is the identity operator.

Along with the operators \hat{B}_i , whose anticommutation relations are satisfied by the three Pauli *spin matrix representations* $B_i = \sigma_i$ (the ‘hat’ will be omitted in denoting matrix representations of operators), two additional operators \hat{A} and \hat{C} have been introduced. We see in equations (26) that the matrix representations of both operators must anticommute with each of the B_i matrix representations, but unlike the Pauli matrices, they neither anticommute with each other nor does each square to the identity matrix (both are in fact singular). Noticing that anticommutation with the B_i is maintained when the representations A and C are each replaced by any linear combination of the two, we seek such linear combinations, B_4 and B_5 , having the same anticommutation relations as the Pauli matrices in order to take advantage of their well documented properties. It is not difficult to demonstrate that they must have the general forms, up to a choice of sign on i ,

$$B_4 = \frac{1}{2}\left(\frac{A}{a} + \frac{C}{b}\right), \quad B_5 = -\frac{i}{2}\left(\frac{A}{a} - \frac{C}{b}\right), \tag{27}$$

provided the nonzero dimensionless constants a and b satisfy

$$2ab = -1. \tag{28}$$

That is, when A , C , and the B_i satisfy equations (26), then equations (27) and (28) insure that the five matrices B_μ , $\mu = 1, \dots, 5$, satisfy the anticommutation relations

$$B_\mu B_\nu + B_\nu B_\mu = 2\delta_{\mu\nu} I, \quad \mu, \nu = 1, \dots, 5. \tag{29}$$

Our situation is reminiscent of Dirac’s derivation of his relativistic wave equation, where he found that the three 2×2 Pauli matrices were not sufficient to satisfy his anticommutation relations. It was necessary to generalize the three Pauli matrices to four 4×4 matrices α_p , $p = 1, \dots, 4$, for which

$$\alpha_p \alpha_q + \alpha_q \alpha_p = 2\delta_{pq} I, \quad p, q = 1, \dots, 4, \tag{30}$$

where I is the 4×4 identity matrix (a demonstration of the necessity for 4×4 matrices is given in, for example [20], p 8). We see now, in light of equation (29), that our proposed linearization of the Schrödinger equation can be completed by finding *five* complex 4×4 matrices satisfying Dirac’s algebra. The definitions of equation (27) are then easily inverted to get expressions for the original matrices A and C of equation (24), viz.,

$$A = a(B_4 + iB_5), \quad C = b(B_4 - iB_5). \tag{31}$$

Fortunately, if four initial matrices satisfying equation (30) are *known*, then the matrix $B_5 = B_1 B_2 B_3 B_4$ is the desired fifth matrix, i.e. a calculation using the properties of the B_p matrices, $p = 1, \dots, 4$, shows that $B_5^2 = I$, and $B_5 B_p = -B_p B_5$ for each p taken separately. For example, the demonstration that $B_5^2 = I$ can be accomplished in three steps: (1) focusing on the first factor of B_5 in the square, use the anticommutation relations to move B_1 three places to the right, and replace $B_1^2 = I$, (2) move B_4 two places to the right, and replace $B_4^2 = I$, and (3) commute B_3 with B_2 , replacing $B_2^2 = B_3^2 = I$ in the result. Each odd

commutation changes the sign on the product, so we have

$$\begin{aligned}
 B_5^2 &= (B_1 B_2 B_3 B_4)(B_1 B_2 B_3 B_4), \\
 &= -B_2 B_3 B_4 B_1^2 B_2 B_3 B_4, \\
 &= -B_2 B_3 B_4 B_2 B_3 B_4, \\
 &= -B_2 B_3 B_2 B_3 B_4^2, \\
 &= -B_2 B_3 B_2 B_3, \\
 &= B_2^2 B_3^2 = I.
 \end{aligned}$$

By similar manipulations, one finds that $B_5 B_p = -B_p B_5$ for each p taken separately.

We emphasize that our matrices B_μ , $\mu = 1, \dots, 5$ are *not* the same as those introduced by either Lévy-Leblond in [25], Greiner in [2] (chapter 13), Hladik [26], or Du Toit [27]. In all four references their matrices are intimately tied to primed counterparts that we have not introduced (and, incidentally, are different in each reference), and products with their primed counterparts *do not* satisfy the anticommutation relations (29) of the Dirac algebra.

4.2. Constructing representations of the Dirac algebra

The first four of our five matrices could be chosen to be Dirac's original matrices, i.e. $B_p = \alpha_p$, $p = 1, \dots, 4$. They can be constructed as Kronecker products of Pauli matrices [28–30], where the Kronecker product of two 2×2 matrices M and N is a 4×4 matrix defined by

$$\begin{aligned}
 M \otimes N &= \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \otimes N = \begin{pmatrix} m_{11}N & m_{12}N \\ m_{21}N & m_{22}N \end{pmatrix}, \\
 &= \begin{pmatrix} m_{11}n_{11} & m_{11}n_{12} & m_{12}n_{11} & m_{12}n_{12} \\ m_{11}n_{21} & m_{11}n_{22} & m_{12}n_{21} & m_{12}n_{22} \\ m_{21}n_{11} & m_{21}n_{12} & m_{22}n_{11} & m_{22}n_{12} \\ m_{21}n_{21} & m_{21}n_{22} & m_{22}n_{21} & m_{22}n_{22} \end{pmatrix}.
 \end{aligned} \tag{32}$$

Recalling the Pauli matrices in their standard representation,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{33}$$

Dirac's original matrices are the following Kronecker products ($i = 1, 2, 3$):

$$\alpha_i = \sigma_i \otimes \sigma_i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \sigma_i = \begin{pmatrix} \mathbb{O} & \sigma_i \\ \sigma_i & \mathbb{O} \end{pmatrix}, \tag{34a}$$

$$\beta = \alpha_4 = \sigma_3 \otimes \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \mathbf{1} = \begin{pmatrix} \mathbf{1} & \mathbb{O} \\ \mathbb{O} & -\mathbf{1} \end{pmatrix}, \tag{34b}$$

where \mathbb{O} is the 2×2 zero matrix, and $\mathbf{1}$ is the 2×2 identity matrix introduced earlier. A fifth matrix of the Dirac algebra is then given by

$$B_5 = \alpha_1 \alpha_2 \alpha_3 \alpha_4 = \begin{pmatrix} \mathbb{O} & -i\mathbf{1} \\ i\mathbf{1} & \mathbb{O} \end{pmatrix}, \tag{35}$$

using the Pauli spin matrix identity $\sigma_1 \sigma_2 \sigma_3 = i\mathbf{1}$. One easily verifies that any pair of the five matrices $B_i = \alpha_i$, $i = 1, 2, 3$, $B_4 = \alpha_4$, and B_5 satisfies equation (29).

A more convenient set of matrices, however, is obtained from the following Kronecker products ($i = 1, 2, 3$):

$$B_i = \sigma_3 \otimes \sigma_i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \sigma_i = \begin{pmatrix} \sigma_i & \mathbb{0} \\ \mathbb{0} & -\sigma_i \end{pmatrix}, \quad (36a)$$

$$B_4 = \sigma_1 \otimes \mathbf{1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \mathbf{1} = \begin{pmatrix} \mathbb{0} & \mathbf{1} \\ \mathbf{1} & \mathbb{0} \end{pmatrix}, \quad (36b)$$

in which case the fifth matrix is

$$B_5 = B_1 B_2 B_3 B_4 = \begin{pmatrix} \mathbb{0} & i\mathbf{1} \\ -i\mathbf{1} & \mathbb{0} \end{pmatrix}. \quad (36c)$$

These matrices also satisfy equations (29), and with their choice the matrices A and C of equations (31) take the simple forms

$$A = 2a \begin{pmatrix} \mathbb{0} & \mathbb{0} \\ \mathbf{1} & \mathbb{0} \end{pmatrix}, \quad C = 2b \begin{pmatrix} \mathbb{0} & \mathbf{1} \\ \mathbb{0} & \mathbb{0} \end{pmatrix}. \quad (37)$$

4.3. The Lévy-Leblond and Pauli equations

In the representations of equations (36a) and (37) the linearized Schrödinger equation (24) takes the matrix form

$$\begin{aligned} & \left(\frac{A}{c} \hat{E} + B_i \hat{p}_j \delta_{ij} + mc C \right) \Psi \\ &= \left[\frac{2a}{c} \begin{pmatrix} \mathbb{0} & \mathbb{0} \\ \mathbf{1} & \mathbb{0} \end{pmatrix} \hat{E} + \begin{pmatrix} \sigma_i & \mathbb{0} \\ \mathbb{0} & -\sigma_i \end{pmatrix} \hat{p}_i + 2bmc \begin{pmatrix} \mathbb{0} & \mathbf{1} \\ \mathbb{0} & \mathbb{0} \end{pmatrix} \right] \Psi, \\ &= \frac{1}{c} \begin{pmatrix} c \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} & 2bmc^2 \mathbf{1} \\ 2a\mathbf{1}\hat{E} & -c \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \end{pmatrix} \Psi = \mathbf{0}, \end{aligned} \quad (38)$$

where we have omitted the position and time arguments of Ψ for convenience, and note that Ψ must now be a bispinor, i.e. a 4-element column matrix of complex functions ($\mathbf{0}$ is the zero bispinor). Equation (38) is the *Lévy-Leblond equation* for a free particle of mass m . Various forms of it can be found in the literature, depending on the choices made for a and b subject to condition (28). We write the bispinor Ψ as

$$\Psi = \begin{pmatrix} \psi \\ \chi \end{pmatrix}, \quad (39)$$

where ψ and χ are two-component spinors (sometimes referred to as semispinors), to obtain the component equations

$$\begin{aligned} \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \psi + 2bmc\mathbf{1}\chi &= \mathbf{0} \\ -2a\mathbf{1}\hat{E} \psi + c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \chi &= \mathbf{0}. \end{aligned}$$

When we choose $a = -1/2$, hence $b = 1$, these reduce to Lévy-Leblond's equations (25) of [25], up to missing factors of c required for correct SI units. Both Greiner's ([2], equation (13.28)), and Hladik's [26], versions of the Lévy-Leblond equations correspond to the choice $a = i/2$, hence $b = i$, while Du Toit's equations [27] correspond to $a = 1/2$, hence $b = -1$. In all that follows, we adhere to Lévy-Leblond's original equations (providing the missing factors of c), hence

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \psi + 2mc\mathbf{1}\chi = \mathbf{0}, \quad (40a)$$

$$\mathbf{1}\hat{E} \psi + c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \chi = \mathbf{0}, \quad (40b)$$

in agreement with the non-relativistic limit of the Dirac equations if c is identified with the speed of light.

If the particle has a charge q and is moving in an external electromagnetic field derived from a vector potential \mathbf{A} and scalar potential ϕ , then equations (40) are modified to include the interaction using the minimal coupling prescription, as was done in equation (6):

$$\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})\psi + 2mc\mathbb{1}\chi = \mathbf{0}, \quad (41a)$$

$$c\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})\chi + (\hat{E} - q\phi)\mathbb{1}\psi = \mathbf{0}. \quad (41b)$$

Eliminating χ between the two equations by multiplying equation (41b) by $2m$, then using equation (41a) to replace $2mc\chi$ in the result, we obtain a wave equation for the spinor ψ :

$$[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})]^2 \psi - 2m(\hat{E} - q\phi)\mathbb{1}\psi = \mathbf{0}, \quad (42)$$

the factors of c obligingly falling out (showing once again that special relativity is not involved). Dividing both sides by $2m$, we see that this is precisely the minimally coupled Schrödinger–Pauli equation, equation (6), from which we have shown that Pauli’s equation (1) follows directly. Thus, as emphasized by Lévy-Leblond in [25], the linearized *non-relativistic* (Galilean covariant) Schrödinger equation together with minimal coupling predicts the correct value for the intrinsic magnetic moment of a spin-1/2 particle: spin is *not* an intrinsically relativistic phenomena.

5. Derivation of the spin current density from the Lévy-Leblond equation

It seems not well known, although Lévy-Leblond published it over 40 years ago, that his equation predicts unambiguously the form of the spin current showing that, like the intrinsic spin itself, it is an inherently non-relativistic phenomena.

To demonstrate this we replace the energy operator in the components, equations (41), of the Lévy-Leblond equation, yielding

$$\frac{\partial\psi}{\partial t} = \frac{ic}{\hbar}\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})\chi - \frac{i}{\hbar}q\phi\mathbb{1}\psi, \quad (43a)$$

$$\chi = -\frac{1}{2mc}\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})\psi, \quad (43b)$$

obvious generalizations of the free particle equations introduced as an ansatz in section 3. In this form we see that χ plays an auxiliary role, as it is given in terms of the component ψ through the constraint (43b). On the other hand, equation (43a) is a true dynamical equation describing the time evolution of the state Ψ via its independent component ψ . Defining the probability density again by $\rho = \psi^\dagger\psi$, we have

$$\frac{\partial\rho}{\partial t} = \psi^\dagger\frac{\partial\psi}{\partial t} + \frac{\partial\psi^\dagger}{\partial t}\psi. \quad (44)$$

The adjoint of (43a) is

$$\frac{\partial\psi^\dagger}{\partial t} = -\frac{ic}{\hbar}[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A})\chi]^\dagger + \frac{i}{\hbar}q\phi\psi^\dagger\mathbb{1}, \quad (45)$$

hence from equations (43a), (45), and (44):

$$\begin{aligned}\frac{1}{c} \frac{\partial \rho}{\partial t} &= \frac{i}{\hbar} \{ \psi^\dagger [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \chi] - [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \chi]^\dagger \psi \}, \\ &= \psi^\dagger (\boldsymbol{\sigma} \cdot \nabla \chi) + (\boldsymbol{\sigma} \cdot \nabla \chi)^\dagger \psi - \frac{iq}{\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi] \\ &\quad - [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi]^\dagger \psi \},\end{aligned}$$

the terms involving $q\phi$ canceling out. The first two terms can be rewritten in terms of a divergence, to obtain

$$\begin{aligned}\frac{1}{c} \frac{\partial \rho}{\partial t} &= \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi) - (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \chi - \chi^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \\ &\quad - \frac{iq}{\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi] - [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi]^\dagger \psi \}.\end{aligned}\quad (46)$$

From equation (43b) and its adjoint, however, we have

$$\begin{aligned}\chi &= -\frac{1}{2mc} [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi], \\ \chi^\dagger &= -\frac{1}{2mc} [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi]^\dagger,\end{aligned}$$

which, when substituted in the last two terms of equation (46) involving \mathbf{A} yield

$$\begin{aligned}-\frac{iq}{\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi] - [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi]^\dagger \psi \} \\ &= \frac{iq}{2mc\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) \boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi] \\ &\quad - [(\boldsymbol{\sigma} \cdot \mathbf{A}) \boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi]^\dagger \psi \}, \\ &= \frac{q}{2mc} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)] + [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)]^\dagger \psi \} \\ &\quad - \frac{iq^2}{2mc\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \mathbf{A} \psi)] \\ &\quad - [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \mathbf{A} \psi)]^\dagger \psi \}.\end{aligned}$$

Here and in what follows we will make frequent use of the fact that $\boldsymbol{\sigma} \cdot \mathbf{A}$ is Hermitian, in which case the last two terms vanish, leaving

$$\begin{aligned}-\frac{iq}{\hbar} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi] - [(\boldsymbol{\sigma} \cdot \mathbf{A}) \chi]^\dagger \psi \} \\ &= \frac{q}{2mc} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)] + [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)]^\dagger \psi \}.\end{aligned}$$

Substituting this result in equation (46) then gives

$$\begin{aligned}\frac{1}{c} \frac{\partial \rho}{\partial t} &= \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi) - [(\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \chi \\ &\quad + \chi^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi)] + \frac{q}{2mc} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)] \\ &\quad + [(\boldsymbol{\sigma} \cdot \mathbf{A}) (\boldsymbol{\sigma} \cdot \nabla \psi)]^\dagger \psi \}.\end{aligned}\quad (47)$$

We next focus attention on the second term, in brackets, of equation (47). Substituting for χ and χ^\dagger in this term yields

$$\begin{aligned}
& (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \chi + \chi^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \\
&= -\frac{1}{2mc} \{ (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi] \\
&\quad + [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi]^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \} \\
&= \frac{i\hbar}{2mc} \{ (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) - (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \} \\
&\quad + \frac{q}{2mc} \{ (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger (\boldsymbol{\sigma} \cdot \mathbf{A} \psi) \\
&\quad + (\boldsymbol{\sigma} \cdot \mathbf{A} \psi)^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \}.
\end{aligned}$$

The first two terms on the right cancel, so we can substitute the last two in equation (47) to obtain an expression with two additional terms in braces:

$$\begin{aligned}
\frac{1}{c} \frac{\partial \rho}{\partial t} &= \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi) + \frac{q}{2mc} \{ \psi^\dagger [(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \nabla \psi)] \\
&\quad + [(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger] \psi - (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger (\boldsymbol{\sigma} \cdot \mathbf{A} \psi) \\
&\quad - (\boldsymbol{\sigma} \cdot \mathbf{A} \psi)^\dagger (\boldsymbol{\sigma} \cdot \nabla \psi) \},
\end{aligned}$$

and again, since $\boldsymbol{\sigma} \cdot \mathbf{A}$ is Hermitian, the terms involving \mathbf{A} vanish, leaving simply

$$\frac{1}{c} \frac{\partial \rho}{\partial t} = \nabla \cdot (\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi),$$

hence the current density can be identified as

$$\mathbf{J} = -c(\psi^\dagger \boldsymbol{\sigma} \chi + \chi^\dagger \boldsymbol{\sigma} \psi). \quad (48)$$

Substituting once more for χ and χ^\dagger , we obtain from equation (48):

$$\begin{aligned}
\mathbf{J} &= \frac{1}{2m} \{ \psi^\dagger \boldsymbol{\sigma} [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi] + [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}) \psi]^\dagger \boldsymbol{\sigma} \psi \} \\
&= -\frac{i\hbar}{2m} [\psi^\dagger \boldsymbol{\sigma} (\boldsymbol{\sigma} \cdot \nabla \psi) - (\boldsymbol{\sigma} \cdot \nabla \psi)^\dagger \boldsymbol{\sigma} \psi] \\
&\quad - \frac{q}{2m} [\psi^\dagger \boldsymbol{\sigma} (\boldsymbol{\sigma} \cdot \mathbf{A} \psi) + (\boldsymbol{\sigma} \cdot \mathbf{A} \psi)^\dagger \boldsymbol{\sigma} \psi].
\end{aligned}$$

From the derivation leading to equations (19) and (20), the first two terms in braces can be replaced to yield

$$\begin{aligned}
\mathbf{J} &= -\frac{i\hbar}{2m} [\psi^\dagger \nabla \psi - (\nabla \psi)^\dagger \psi] + \frac{\hbar}{2m} \nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi) \\
&\quad - \frac{q}{2m} [\psi^\dagger \boldsymbol{\sigma} (\boldsymbol{\sigma} \cdot \mathbf{A} \psi) + \psi^\dagger (\boldsymbol{\sigma} \cdot \mathbf{A}) \boldsymbol{\sigma} \psi].
\end{aligned}$$

The i th component of the terms involving \mathbf{A} can be written, making use of the Pauli matrix identity $\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij}$, very simply as

$$\begin{aligned}
\psi^\dagger \sigma_i (\sigma_j A_j \psi) + \psi^\dagger (\sigma_j A_j) \sigma_i \psi &= \psi^\dagger (\sigma_i \sigma_j + \sigma_j \sigma_i) A_j \psi \\
&= 2\psi^\dagger \delta_{ij} A_j \psi = 2A_i \psi^\dagger \psi,
\end{aligned}$$

which, when substituted in the previous equation, yields the correct non-relativistic current density for a spin-1/2 particle interacting with an external electromagnetic field:

$$\begin{aligned} \mathbf{J} = & -\frac{i\hbar}{2m} [\psi^\dagger \nabla \psi - (\nabla \psi)^\dagger \psi] + \frac{\hbar}{2m} \nabla \times (\psi^\dagger \boldsymbol{\sigma} \psi) \\ & - \frac{q}{m} \mathbf{A} \psi^\dagger \psi, \end{aligned} \quad (49)$$

verifying equations (10) and (11).

6. Does the spin current have measurable consequences?

Although the spin current has no effect in the *probability conservation equation*, it has been convincingly demonstrated to be an inherent feature of the *probability current density* of a spin-1/2 particle described by the Schrödinger–Pauli equation. If the particle carries a charge q , then it should contribute to the *charge current density*, hence be measurable in experiments sensitive to the effects of such currents. After all, paraphrasing Nowakowski in [6], electric current is a physical observable.

In [6] an interesting experiment is proposed and analyzed to demonstrate the relevance of the spin current. It is determined that in such an experiment there would be a contribution to the electric current density ‘which is of purely quantum mechanical origin and which can be traced back to the ‘spin term’ ...’. The author further speculates that ‘this term might play a role in problems concerning conductivity in solids.’

Du Toit, in his honors thesis [27], discusses a simpler experiment, an electron in a homogeneous magnetic field (the Landau problem). His analysis is rather inconclusive, the main effect being described as a ‘swirl’ of charge current creating a current loop/dipole moment that could possibly interact with the field to produce a torque on the dipole, presumably observable as some sort of precessional motion.

A field of physics where the current density plays a *critical* role is Bohmian mechanics (or de Broglie–Bohm theory; see [31]–[33] for excellent general accounts). Well-defined concepts of point particles and trajectories are fundamental features of this theory, where particle motion takes place along trajectories that are integral curves of a velocity field proportional to the current density ([33], chapter 10). Thus, a particle’s equations of motion are directly determined by the current density itself, hence the spin term can directly influence its motion. In recent work [34, 35] directed toward the prediction of arrival times in time-of-flight experiments, the possibility of observing rather remarkable effects of the spin term is discussed. In these papers, time-of-flight experiments are proposed and carefully analyzed for electrons prepared in various initial states, moving in a cylindrical waveguide. Analytical and numerical methods are used to obtain predictions of their arrival time distributions at a detector. For a particular initial state, with spin vector oriented perpendicular to the waveguide axis, the predicted distributions show a striking result, a cutoff, or maximum arrival time, after which no further electrons would be detected. Such an experiment, or one similar to it, seems to be well within the scope of present day technology. The results would be a welcome test of the ability of Bohmian mechanics to predict arrival times of spin-1/2 particles, given that a standard quantum mechanical prediction of the same is still ambiguous [36].

Finally, as further examples of the possible impact of the spin term, we mention recent publications [37, 38] applying the Dirac theory to experiments with relativistic electron beams. In these, a velocity field of probability flow is defined by precisely the same equation used in Bohmian mechanics to describe particle trajectories, i.e. proportional to the current density, and used to describe results of such experiments (or suggested ones). As in the non-relativistic case, the spin term makes a definite contribution that cannot be ignored.

7. Concluding remarks

After reviewing previous work, our goals here were two-fold. First, to emphasize that although the Pauli equation can be *constructed* from the Schrödinger equation by judicious use of spin operator identities, thus demonstrating the non-relativistic nature of the electron spin, Lévy-Leblond pointed the way to a direct *derivation* of the Pauli equation via a linearization of the Schrödinger equation, from which the non-relativistic nature of the electron spin naturally follows. Second, to direct attention to the work of Shikakhwa *et al* [8], who we have argued showed conclusively and unambiguously that the spin term of the probability current density is a non-relativistic phenomenon derivable from the Schrödinger–Pauli equation, contrary to what is sometimes claimed in the literature [6]. Furthermore, we emphasized the importance of the Lévy-Leblond theory in this regard, showing that the spin term is directly derivable from his equation, which is properly viewed as a precursor of the Pauli equation.

We conclude by cautioning readers about the casual use of the ubiquitous minimal coupling prescription in this and other work, invoking electromagnetic potentials that are inherently relativistic (since they originate in the Lorentz covariant Maxwell equations), and using them in a non-relativistic quantum theory involving the Schrödinger, Pauli, or Lévy-Leblond equations. Constraints on the allowable electromagnetic fields as a result of requiring Galilean covariance of both the quantum theory and electromagnetism were first discussed by Lévy-Leblond in [25], p 305. His results have been sharpened and developed considerably in later work [39–44], but a discussion of these important articles is beyond the scope of the present paper. A review of this work is highly recommended to anyone applying non-relativistic quantum theories to the analysis of experimental results involving electromagnetic interactions.

Acknowledgments

My thanks to Siddhant Das for his encouragement in pursuing a better understanding of the Pauli equation and its spin current density, and for critically reviewing the manuscript. In addition, the comments of reviewers were invaluable, especially in clarifying the discussion of section 2, and correcting typographical errors in section 3. I am grateful to Detlef Dürr for his interest in this work, and for his collegial grace in making me feel welcome as a long-distance collaborator on research efforts at the Mathematisches Institut, Ludwig-Maximilians-Universität München.

ORCID iDs

James M Wilkes  <https://orcid.org/0000-0003-1207-4122>

References

- [1] Levin F S 2002 *An Introduction to Quantum Theory* (New York: Cambridge) p 499
- [2] Greiner W 2001 *Quantum Mechanics: An Introduction* 4 edn (Berlin: Springer) pp 339–40
- [3] Ikenberry E 1962 *Quantum Mechanics for Mathematicians and Physicists* (New York: Oxford) pp 241–2
- [4] Galindo A and Sanchez Del Rio C 1961 Intrinsic magnetic moment as a non-relativistic phenomenon *Am. J. Phys.* **29** 582–4

- [5] Parker G W 1984 Spin current density and the hyperfine interaction in hydrogen *Am. J. Phys.* **52** 36–9
- [6] Nowakowski M 1999 The quantum mechanical current of the Pauli equation *Am. J. Phys.* **67** 916–9
- [7] Mita K Virtual probability current associated with the spin *Am. J. Phys.* **68** 259–64
- [8] Shikakhwa M S, Turgut S and Pak N K 1999 Derivation of the magnetization current from the non-relativistic Pauli equation: a comment on ‘The quantum mechanical current of the Pauli equation’ by Marek Nowakowski [*Am. J. Phys.* **67** (10), 916 (1999)] *Am. J. Phys.* **79** 1177–9
- [9] Hodge W B, Migirditch S V and Kerr W C 2014 Electron spin and probability current density in quantum mechanics *Am. J. Phys.* **82** 681–90
- [10] Pauli W Jr. 1927 Zur quantenmechanik des magnetischen elektrons *Z. Phys.* **43** 601–23
- [11] Goudsmit S and Uhlenbeck G E 1925 Opmerking over de spectra van Waterstof in helium *Physica* **5** 266–70 (<https://repository.aip.org/islandora/object/nbla%3A227343#page/1/>)
- [12] Uhlenbeck G E and Goudsmit S 1925 Ersetzung der hypothese vom unmechanischen zwang durch eine forderung bezüglich des innern verhaltens jedes einzelnen electrons *Naturwissenschaften* **13** 953–4
- [13] Dirac P A M 1928 The quantum theory of the electron *Proc. R. Soc. A* **117** 610–24
- [14] Lévy-Leblond J-M 1974 The pedagogical role and epistemological significance of group theory in quantum mechanics *Riv. Nuovo Cimento* **4** 99–143
- [15] Feynman R P 1961 *Quantum Electrodynamics* vol 37 (New York: Addison-Wesley) pp 6–40
- [16] Sakurai J J 1967 *Advanced Quantum Mechanics* (Reading, MA: Addison-Wesley) pp 78–80
- [17] Adler R J and Martin R A 1992 The electron g factor and factorization of the Pauli equation *Am. J. Phys.* **60** 837–9
- [18] Griffiths D J 1999 *Introduction to Electrodynamics* 3rd edn (Upper Saddle River, NJ: Prentice Hall) pp 263–8
- [19] Russakoff G 1970 A derivation of the macroscopic Maxwell equations *Am. J. Phys.* **38** 1188–95
- [20] Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill) pp 10–3
- [21] Landau L D and Lifshitz E M 1965 *Quantum Mechanics: Non-Relativistic Theory* 2nd edn (Oxford: Pergamon) pp 435–7
- [22] Holland P 1999 Uniqueness of paths in quantum mechanics *Phys. Rev. A* **60** 4326–30
- [23] Holland P 2003 Uniqueness of conserved currents in quantum mechanics *Ann. Phys.* **12** 446–62
- [24] Holland P and Philippidis C 2003 Implications of Lorentz covariance for the guidance equation in two-slit quantum interference *Phys. Rev. A* **67** 062105-1–15
- [25] Lévy-Leblond J-M 1967 Non-relativistic particles and wave equations *Commun. Math. Phys.* **6** 286–311
- [26] Hladik J 1999 *Spinors in Physics* (New York: Springer) pp 100–5
- [27] Toit P D 2001 Spin current: the contribution of spin to the probability current for non-relativistic particles *Honors Thesis* Brigham Young University (<https://physics.byu.edu/thesis/archive/2001/63>)
- [28] Arfken G 1995 *Mathematical Methods for Physicists* (New York: Academic) pp 179–211
- [29] Snygg J 1997 *Clifford Algebra: A Computational Tool for Physicists* (New York: Oxford) p 292
- [30] Goodmanson D M 1996 A graphical representation of the Dirac algebra *Am. J. Phys.* **64** 870–80
- [31] Dürr D and Teufel S 2009 *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Berlin: Springer)
- [32] Holland P R 2000 *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics* (New York: Cambridge)
- [33] Bohm D and Hiley B J 1993 *The Undivided Universe: An Ontological Interpretation of Quantum Theory* (London and New York: Routledge)
- [34] Das S and Dürr D 2019 Arrival time distributions of spin-1/2 particles *Sci. Rep.* **9** 2242–50
- [35] Das S, Nöth M and Dürr D 2019 Exotic Bohmian arrival times of spin-1/2 particles: an analytical treatment *Phys. Rev. A* **99** 052124-1–13
- [36] Muga J G and Leavens C R 2000 Arrival time in quantum mechanics *Phys. Rep.* **338** 353–438
- [37] Bialynicki-Birula I and Bialynicka-Birula Z 2017 Relativistic electron wave packets carrying angular momentum *Phys. Rev. Lett.* **118** 114801-1–5
- [38] Bialynicki-Birula I and Bialynicka-Birula Z 2019 Twisted localized solutions of the Dirac equation: hopfionlike states of relativistic electrons *Phys. Rev. A* **100** 012108-1–5
- [39] Bellac M L and Lévy-Leblond J-M 1973 Galilean electromagnetism *Nuovo Cimento B* **14** 217–33

- [40] Brown H R and Holland P R 1999 The Galilean covariance of quantum mechanics in the case of external fields *Am. J. Phys.* **67** 204–14
- [41] Holland P R and Brown H R 2003 The non-relativistic limits of the Maxwell and Dirac equations: the role of galilean and gauge invariance *Stud. Hist. Phil. Mod. Phys.* **34** 161–87
- [42] de Montigny M and Rousseaux G 2006 On the electrodynamics of moving bodies at low velocities *Eur. J. Phys.* **27** 755–68
- [43] de Montigny M and Rousseaux G 2007 On some applications of Galilean electrodynamics of moving bodies *Am. J. Phys.* **75** 984–92
- [44] Manfredi G 2013 Non-relativistic limits of Maxwell’s equations *Eur. J. Phys.* **34** 859–71