Foundations of Physics, Vol. 28, No. 3, 1998

Probability Backflow for a Dirac Particle

G. F. Melloy[1] **and A. J. Bracken**[1]

Received August 15, 1997

The phenomenon of probability backflow, previously quantified for a free nonrelativistic particle, is considered for a free particle obeying Dirac's equation. It is shown that probability backflow can occur in the opposite direction to the momentum; that is to say, there exist positive-energy states in which the particle certainly has a positive momentum in a given direction, but for which the component of the probability flux vector in that direction is negative. It is shown that the maximum possible amount of probability that can flow "backwards," over a given time interval of duration T, depends on the dimensionless parameter $\varepsilon = \sqrt{4\hbar/mc^2 T}$, *where m is the mass of the particle and c is the speed of light. At* $\varepsilon = 0$, *the nonrelativistic value of approximately 0.039 for this maximum is recovered. Numerical studies suggest that the maximum decreases monotonically as* ε *increases from 0, and show that it depends on the size of m,* $\hbar$, *and T, unlike the nonrelativistic case.*

1. INTRODUCTION

Asim Barut was intrigued by the remarkable structure of Dirac's equation for the electron: The meaning of the *Zitterbewegung*,[1] the structural role of the group $SO(4, 2)$,[2] and the interpretation of different concepts of localization[3] for the electron were among topics he helped to illuminate with characteristically novel insights. One aspect which he repeatedly emphasized in discussions was the independence of the concepts of velocity and momentum for the Dirac electron. In the Heisenberg picture, the equation of motion for the Dirac coordinate $\mathbf{x}$ is of third order: with H the Dirac Hamiltonian, we have in the case of the free particle[1]

$$i\hbar\dddot{\mathbf{x}} = -2H\ddot{\mathbf{x}} \tag{1}$$

allowing $\mathbf{x}$, $\dot{\mathbf{x}}$, and $\ddot{\mathbf{x}}$ (or $\mathbf{x}$, $\dot{\mathbf{x}}$ and the momentum $\mathbf{p}$) to be independent operator functions of time.

[1] Department of Mathematics, The university of Queensland, Brisbane 4072, Australia.

It was shown recently[4] that the quantum mechanical description of a free nonrelativistic particle allows for the existence of states in which the probability of finding the particle to have a *negative* x-coordinate increases with time even though the states certainly have *positive* x-component of momentum (velocity). The possibility of such a remarkable and counter-intuitive nonclassical effect seems to have been indicated first by Allcock.[5] The maximum possible amount of probability backflow in any given time interval T, taken over all possible initial positive-momentum states, was found[4] to have the value

$$\Delta_{max} \approx 0.039 \tag{2}$$

This value is independent of the size of T, the mass m of the particle, and most remarkably, of Planck's constant $\hbar$.

As the description of the nonrelativistic behavior of a particle must correspond to a limiting case of its relativistic description, it is to be expected that the phenomenon of probability backflow will occur in the relativistic case. However, in the case of the free Dirac particle, it is not immediately clear what are the appropriate states to consider, because of the nontrivial distinction between velocity and momentum. Furthermore, the concept of localization of such a particle is contentious,[3] and the phenomenon of *Zitterbewegung* [1] in particular raises questions as to the expected behavior of a Dirac particle in regard to probability backflow.

The following analysis shows that probability flow can indeed take place against the direction of the *momentum* of a free Dirac particle. It transpires that the maximum possible cumulative amount of probability flowing "in the wrong direction" over a time interval of duration T depends on the dimensionless parameter $\varepsilon = \sqrt{4\hbar/mc^2 T}$, involving the speed of light c, as might be anticipated.[4] Numerical analysis suggests that this maximum possible backflow decreases monotonically towards zero with increasing ε, and so depends on the sizes of T, $\hbar$, and m, unlike the non-relativistic case. For any fixed values of these three quantities, the nonrelativistic result is recovered as $c \to \infty$.

2. ANALYSIS

It is sufficient to consider Dirac's equation in one dimension to demonstrate the result. The wavefunction in this case has two complex-valued components:

$$\Psi(x, t) = \begin{pmatrix} \psi_1(x, t) \\ \psi_2(x, t) \end{pmatrix} \tag{3}$$

If the rest mass of the particle is m, the Hamiltonian can be expressed in terms of Pauli spin matrices as

$$H = c\sigma_1 p + \sigma_3 mc^2 \tag{4}$$

where $p = -i\hbar\,\partial/\partial x$, and Dirac's equation

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = H\Psi(x, t) \tag{5}$$

can be expressed as the two equations

$$\frac{\partial \psi_1(x, t)}{c\,\partial t} + \frac{\partial \psi_2(x, t)}{\partial x} = -\frac{imc}{\hbar}\psi_1(x, t) \tag{6}$$

$$\frac{\partial \psi_2(x, t)}{c\,\partial t} + \frac{\partial \psi_1(x, t)}{\partial x} = \frac{imc}{\hbar}\psi_2(x, t) \tag{7}$$

The expressions for the probability density $\rho = \Psi^\dagger\Psi$ and probability flux vector $j = c\Psi^\dagger\sigma_1\Psi$ take the forms

$$\rho(x, t) = \psi_1^*(x, t)\,\psi_1(x, t) + \psi_2^*(x, t)\,\psi_2(x, t) \tag{8}$$

and

$$j(x, t) = c(\psi_1^*(x, t)\,\psi_2(x, t) + \psi_2^*(x, t)\,\psi_1(x, t)) \tag{9}$$

respectively.

We consider positive momentum solutions to Dirac's equation, of the form

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_0^\infty e^{-iHt/\hbar} e^{ixp/\hbar} \Phi(p)\, dp \tag{10}$$

where

$$\Phi(p) = \begin{pmatrix} \phi_1(p) \\ \phi_2(p) \end{pmatrix} \tag{11}$$

is the initial state vector in the momentum representation. The integration limits in (10) ensure that only positive momentum components are considered. Furthermore, for physically meaningful solutions, we are restricted to positive energy states. Thus we have

$$[cp\sigma_1 + mc^2\sigma_3]\,\Phi(p) = E(p)\,\Phi(p) = \sqrt{c^2p^2 + m^2c^4}\,\Phi(p) \tag{12}$$

so that

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_0^\infty e^{ixp/\hbar} e^{-iE(p)t/\hbar} \Phi(p)\, dp \tag{13}$$

with the normalization condition

$$\int_0^\infty \Phi^\dagger(p)\, \Phi(p)\, dp = 1 \tag{14}$$

Positive energy solutions, satisfying (12) and (14), have the general form

$$\begin{aligned} \Phi(p) &= f(p) \frac{1}{\sqrt{2E(p)(E(p) + mc^2)}} \begin{pmatrix} mc^2 + E(p) \\ cp \end{pmatrix} \\ &= f(p) \begin{pmatrix} u_1(p) \\ u_2(p) \end{pmatrix} = f(p)\, u(p) \end{aligned} \tag{15}$$

with $f(p)$ an arbitrary complex-valued function on $0 \leqslant p < \infty$ satisfying the normalization condition

$$\int_0^\infty f^*(p)\, f(p)\, dp = 1 \tag{16}$$

From (13) and (15) we can now express the probability flux vector (9) at the point $x = 0$ as

$$\begin{aligned} j(0, t) &\\ &= \frac{c}{2\pi\hbar} \int_0^\infty e^{iE(p)t/\hbar} \phi_1^*(p)\, dp \int_0^\infty e^{-iE(q)t/\hbar} \phi_2(q)\, dq + \text{c.c.} \\ &= \frac{c}{2\pi\hbar} \int_0^\infty \int_0^\infty e^{it(E(p) - E(q))/\hbar} f^*(p)\, f(q) [u_1(p)\, u_2(q) + u_1(q)\, u_2(p)]\, dq\, dp \end{aligned} \tag{17}$$

From an intuitive, "classical" perspective, $j(x, t)$ should be nonnegative at all values of x and t, because the particle certainly has a nonnegative x-component of momentum as a result of (10). In fact however, $j(x, t)$ can be negative, in particular at $x = 0$, and there can therefore be a net backflow of probability from the region $x > 0$ to the region $x < 0$ over a given time interval. For the interval $(0, T)$, this net backflow is given by

$$\begin{aligned} \Delta &= -\int_0^T j(0, t)\, dt \\ &= -\frac{c}{2\pi\hbar}\int_0^\infty \int_0^\infty \frac{e^{iT(E(p)-E(q))/\hbar} - 1}{i(E(p)-E(q))/\hbar} f^*(p) f(q) \\ &\quad \times [u_1(p)\, u_2(q) + u_1(q)\, u_2(p)]\, dq\, dp \\ &= \int_0^\infty \int_0^\infty f^*(p)\, K(p, q)\, f(q)\, dp\, dq \end{aligned} \tag{18}$$

where

$$\begin{aligned} K(p, q) &= \frac{ic}{2\pi} \frac{e^{iT(E(p)-E(q))\hbar} - 1}{E(p)-E(q)} \frac{1}{\sqrt{2E(p)(E(p)+mc^2)}} \frac{1}{\sqrt{2E(q)(E(q)+mc^2)}} \\ &\quad \times \frac{1}{[(mc^2 + E(p))\, cq + (mc^2 + E(q))\, cp]} \end{aligned} \tag{19}$$

Note that this kernel is hermitian and that the singularity at $p = q$ is only apparent. Our object is to maximize Δ; as we shall see, this maximum is indeed positive, implying that $j(0, t)$ can be negative. We have to take into account the normalization constraint (16), and we therefore consider the unconstrained maximum of

$$I(f) = \int_0^\infty \int_0^\infty f^*(p)\, K(p, q)\, f(q)\, dp\, dq - \lambda \int_0^\infty f^*(p)\, f(p)\, dp \tag{20}$$

where λ is a Lagrange multiplier. At any stationary point of I, the Euler–Lagrange equation

$$\int_0^\infty K(p, q)\, f(q)\, dq = \lambda f(p) \tag{21}$$

must hold, and then for such an f, it follows from (18) and (16) that $\Delta = \lambda$. Our problem is then to find the largest positive eigenvalue in (21) of the integral operator with kernel $K(p, q)$.

With the substitutions

$$\varepsilon = \sqrt{\frac{4\hbar}{mc^2 T}} \tag{22}$$

$$p = \sqrt{\frac{4m\hbar}{T}}\, r = mc\varepsilon r \tag{23}$$

$$q = \sqrt{\frac{4m\hbar}{T}}\, s = mc\varepsilon s \tag{24}$$

$$E(p) = \sqrt{c^2p^2 + m^2c^4} = mc^2\sqrt{\varepsilon^2 r^2 + 1} = mc^2 \mathscr{E}(r) \tag{25}$$

the eigenvalue equation becomes

$$\int_0^\infty \frac{i}{4\pi} \frac{(e^{4i(\mathscr{E}(r) - \mathscr{E}(s))/\varepsilon^2} - 1)[r(\mathscr{E}(s) + 1) + s(\mathscr{E}(r) + 1)]}{[(\mathscr{E}(r) - \mathscr{E}(s))/\varepsilon^2]\sqrt{\mathscr{E}(r)\mathscr{E}(s)(\mathscr{E}(r) + 1)(\mathscr{E}(s) + 1)}} f(mc\varepsilon s)\, ds$$
$$= \lambda f(mc\varepsilon r) \tag{26}$$

which reduces to the real equation

$$\int_0^\infty \frac{1}{\pi} \frac{\sin[2(\mathscr{E}(r) - \mathscr{E}(s))/\varepsilon^2][r(\mathscr{E}(s) + 1) + s(\mathscr{E}(r) + 1)]}{[2(\mathscr{E}(r) - \mathscr{E}(s))/\varepsilon^2]\sqrt{\mathscr{E}(r)(\mathscr{E}(r) + 1)\mathscr{E}(s)(\mathscr{E}(s) + 1)}} \varphi(s)\, ds = -\lambda\varphi(r) \tag{27}$$

where

$$\varphi(r) = e^{-2i\mathscr{E}(r)/\varepsilon^2} f(mc\varepsilon r) \tag{28}$$

By letting ε approach zero we recover the nonrelativistic eigenvalue equation

$$\frac{1}{\pi}\int_0^\infty \frac{\sin(r^2 - s^2)}{(r - s)} \varphi(s)\, ds = -\lambda\varphi(r) \tag{29}$$

as derived in our earlier study,[4] and for which we know the largest positive eigenvalue is approximately 0.039.

3. SUMMARY OF RESULTS

An analytical solution of the eigenvalue equation (27), or even of its nonrelativistic counterpart (29), has not been forthcoming, and numerical methods have been used to estimate Δ_{max}. With the range of integration in

(27) approximated by the interval $[0, N\tau]$, divided into N segments of length τ, library subroutines were used to estimate the integral offer each subinterval and then to find the largest eigenvalue of the associated discretized eigenvalue problem. For each of a set of values of the parameter ε, it was found that as $\tau \to 0$ and $N\tau \to \infty$, the maximum value of backflow apparently approached a limit Δ_{max}. The estimates obtained were

$$\begin{aligned} \varepsilon &= 0.01, & \Delta_{max} &= 0.0384 \\ \varepsilon &= 0.1, & \Delta_{max} &= 0.0376 \\ \varepsilon &= 0.5, & \Delta_{max} &= 0.0312 \\ \varepsilon &= 1.0, & \Delta_{max} &= 0.0249 \end{aligned} \tag{30}$$

suggesting that Δ_{max} decreases monotonically as $\varepsilon \to \infty$.

These results may be interpreted in a number of ways. Firstly, they indicate the variation in Δ_{max} as the velocity of light c "varies" for fixed $\hbar$, m, and T. Then $c \to \infty$ corresponds to $\varepsilon \to 0$, where the estimates of Δ_{max} converge to the nonrelativistic value of approximately 0.039. Alternatively, the results show the variation in Δ_{max} for varying values of $\hbar$, m, or T when c is held fixed. Then it is apparent that the greatest value of Δ_{max} corresponds to $T \to \infty$, $m \to \infty$, or $\hbar \to 0$. This last result is surprising, as $\hbar \to 0$ is usually regarded as corresponding to the classical limit, yet the phenomenon of probability backflow is certainly a nonclassical effect.

These results should dispel any suspicion that the emergence of the peculiar phenomenon of probability backflow—or of the mysterious quantum number whose approximate value is 0.039—indicates some inadequacy in the nonrelativistic formulation of quantum mechanics.

4. PHYSICAL INTERPRETATION

The phenomenon of probability backflow was shown in the nonrelativistic case to be a kind of self-interference effect for quantum wavepackets. This was illustrated by considering a (non-normalizable) state which was a superposition of two plane waves.

It is interesting to consider an analogous state for the relativistic case as this throws some light on the nature and physical meaning of the phenomenon.

Consider then the wave function

$$\Psi(x, t) = Au(p_1)\, e^{i\theta_1(x, t)} + Bu(p_2)\, e^{i\theta_2(x, t)} \tag{31}$$

with

$$\theta_n(x, t) = \frac{p_n x - E(p_n)\, t}{\hbar} + \gamma_n \tag{32}$$

and

$$u(p_n) = \begin{pmatrix} u_1(p_n) \\ u_2(p_n) \end{pmatrix}, \qquad n = 1, 2 \tag{33}$$

Here A and B are non-negative constants, p_1 and p_2 are positive momenta, and γ_1 and γ_2 are arbitrary constant phases. This leads to a current distribution

$$\begin{aligned} j(x, t) &= c\Psi^\dagger(x, t)\, \sigma_1\, \Psi(x, t) \\ &= 2c(A^2 u_1(p_1)\, u_2(p_1) + B^2 u_1(p_2)\, u_2(p_2) + AB[u_1(p_1)\, u_2(p_2) \\ &\quad + u_1(p_2)\, u_2(p_1)] \cos[\theta_2(x, t) - \theta_1(x, t)]) \end{aligned} \tag{34}$$

For a single plane wave with positive momentum ($B = 0$) this gives a positive constant, as might be expected for a particle traveling in the $+x$-direction. However, for two such plane waves, so that A and B are both positive, the flux varies with time between an upper value of

$$(Au_1(p_1) + Bu_1(p_2))(Au_2(p_1) + Bu_2(p_2)) \tag{35}$$

and a lower value of

$$(Au_1(p_1) - Bu_1(p_2))(Au_2(p_1) - Bu_2(p_2)) \tag{36}$$

This lower value can be made negative by a suitable choice of A, B, p_1, and p_2. For example, if

$$p_2 = 2p_1 = 2mc \tag{37}$$

we can choose

$$A = \left(\frac{4 + 2\sqrt{2}}{10 + 2\sqrt{5}}\right) \alpha B \tag{38}$$

with

$$\frac{1 + \sqrt{5}}{1 + \sqrt{2}} < \alpha < 2 \tag{39}$$

We then have

$$u_1(p_1) = (1+\sqrt{2})\, u_2(p_1) = \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \tag{40}$$

and

$$u_1(p_2) = (1+\sqrt{5})\, u_2(p_2) = \frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}} \tag{41}$$

and it is easily checked that the value of the expression (36) is indeed negative.

Can the possibility of negative values of j be related to the independence of the velocity operator $\dot{x} = c\sigma_1$ and the momentum operator p? Given $j(x, t)$, we may define a "local mean velocity" $v(x, t)$ by

$$v(x, t) = \frac{j(x, t)}{\rho(x, t)} \tag{42}$$

Because ρ is positive, this local mean velocity is positive or negative according as $j(x, t)$ is positive or negative. We have

$$v(x, t) = \frac{\Psi^\dagger(x, t)\, \dot{x}\, \Psi(x, t)}{\Psi^\dagger(x, t)\, \Psi(x, t)} \tag{43}$$

and since $\dot{x}$ commutes with x, it makes sense to refer to $v(x, t)$ as a "local" expectation value of $\dot{x}$. Note that the overall expectation value of $\dot{x}$ in the state Ψ is

$$\langle \dot{x} \rangle = \int_{-\infty}^{\infty} \Psi^\dagger(x, t)\, \dot{x}\, \Psi(x, t)\, dx = \int_{-\infty}^{\infty} \rho(x, t)\, v(x, t)\, dx \tag{44}$$

As is well known,

$$\langle \dot{x} \rangle (t) = \langle c^2 p/E(p) \rangle \tag{45}$$

which is constant and positive in the cases under consideration. We see that the possibility of negative values of $j(x, t)$ (equivalently of $v(x, t)$) is directly related to the association of negative as well as positive values of $\dot{x}$ with each value of x, even though p is positive, and even though we are dealing with positive energy states, in which the sign of the velocity $\dot{x}$ is indeterminate.

In principle the phenomenon of probability backflow for the Dirac particle is observable; by experiments on many systems prepared in the same state, we can measure the probability mass on the negative x-axis after any elapsed time, and hence the changes of this quantity with time. It would seem important to confirm by experiment the predicted maximum possible amount of probability that can flow "backwards" over a given time interval, since this would help to confirm the description of the particle by Dirac's equation. Even an observation of the phenomenon might be regarded as significant in providing a direct confirmation of the independence of the concepts of velocity and momentum for the Dirac particle, and for this reason alone it is worthy of experimental investigation. In particular, it is remarkable that negative values of $v(x, t)$ might be observed in positive energy states, even though in such states the overall expectation value $\langle \dot{x} \rangle (t)$ is positive.

We should keep in mind that the probability backflow phenomenon can also occur for a nonrelativistic, spinless particle, where there seems to be no possibility to attribute its origin to the independence of the concepts of momentum and velocity. In that case too we can define a velocity $v(x, t)$ in terms of probability density and current density by an equation like (42), and again this velocity will necessarily be in the same direction as the current density. However, the interpretation of such a $v(x, t)$ as a "local mean velocity" is more difficult to justify in this case, and the interpretation of the backflow phenomenon is accordingly more obscure.

ACKNOWLEDGMENT

We wish to thank the referees for their comments on the preliminary version of this paper, which encouraged us to expand on the physical interpretation of the phenomenon.

REFERENCES

1. A. O. Barut and A. J. Bracken, *Phys. Rev. D* **23**, 2454–2463 (1981).
2. A. O. Barut, *Phys. Rev. Lett*. **20**, 893 (1968).
3. A. O. Barut and S. Malin, *Rev. Mod. Phys.* **40**, 632–640 (1968).
4. A. J. Bracken and G. F. Melloy, *J. Phys. A* **27**, 2197–2211 (1994).
5. G. R. Allcock, *Ann. Phys.* **53**, 311–348 (1969).