J. Phys. A: Math. Gen. **29** (1996) 5181–5186. Printed in the UK

A direct derivation of the Dirac equation via quaternion measures

S K Srinivasan† and E C G Sudarshan‡§

† Department of Mathematics, Indian Institute of Technology, Madras - 600 036, India
‡ Center for Particle Physics and Department of Physics, University of Texas, Austin, TX 78712, USA‖
§ School of Physics, University of Hyderabad, Hyderabad - 500134, India

Received 28 February 1996

Abstract. Quaternion measurable processes are introduced and the Dirac equation is derived from the Langevin equation associated with a two-valued process.

1. Introduction

The object of this contribution is to show that the relativistic equation for spin-$\frac{1}{2}$ particles can be obtained by an enlargement of the theory of stochastic processes one step beyond the theory of complex measures. In the conventional formulation of the Schrödinger and the Dirac equations the evolution is in terms of wavefunctions which themselves do not have a probabilistic interpretation. It is indeed possible to dispense with such a starting point and, instead, formulate the problem in terms of a complex (vector) measure. In an earlier contribution [1], we have provided a new derivation of the Schrödinger equation by identifying it as the Fokker–Plank equation corresponding to a complex measurable Markov process. We now generalize the field to quaternions and consider quaternion measurable processes.

2. Quaternion measures and Pauli systems

The quaternion measure is introduced in exactly the same way as the complex measures; a quaternion measure λ for any set $A \in B$ by

$$\lambda(A) = \lambda_0(A) + \hat{i}\lambda_1(A) - \hat{j}\lambda_2(A) + \hat{k}\lambda_3(A) \tag{1}$$

where $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ are complex measures (see, for example, [2]) defined on the measurable space (Ω, B) and $\hat{i}, \hat{j}, \hat{k}$ are the hypercomplex numbers introduced by Hamilton (see [3]). Clearly such a quaternion measure may be seen as a member of the wider class of vector measures dealt with in classical analysis.

A quaternion measurable process is an indexed family of such measures. We can proceed to introduce, systematically, random variables and random processes in the manner outlined by Pitt [4]. The concepts of conditional probabilities can be introduced in exactly the same

† Emeritus.
‖ Permanent address.

0305-4470/96/165181+06\$19.50

way as was done for complex measures [1]; we can define Markov processes and derive the Chapman–Kolmogorov relation.

Within such a framework of quaternion measurable processes, we now consider the Langevin equation

$$\mathrm{d}x^i = v^j \exp(\mathrm{i}\pi z(t))\,\mathrm{d}t \qquad j = 1, 2, 3 \tag{2}$$

where t is the time parameter and $z(t)$ is a two-valued process on 0, 1. We further constrain $z(t)$ to be a Markov chain with rates $\lambda_\pm$ of transition.

Before we proceed to deal with quaternion measurable processes, we note that the very introduction of complex measures has some implications:

(i) a real valued random variable has a complex valued expectation in general;

(ii) if a, the complex expectation value, is the result of a particular choice of a distribution, then there is a minimum value for the spread of a random variable X where for arbitrary $\epsilon > 0$ spread of X is defined by (see [1])

$$\text{spread of } X = |\Pr\{|X - E[X]| > \epsilon\}|.$$

Thus if we introduce a quaternion measure, then the expected value of a real valued random variable is also a quaternion. We illustrate this for the simple case of a random walk problem dealt with earlier [1], which is now in the framework of quaternion measurable processes (QMP). Assume that in the random walk in discrete steps, the steps are given by

$$\begin{aligned}
\Delta r &= \pm 1 && \text{in to the } x\text{-direction with (quaternion) probability } p_{1\pm} + \sigma_1 q_1 \\
&= \pm 1 && \text{in to the } y\text{-direction with probability } p_{2\pm} + \sigma_2 q_2 \\
&= \pm 1 && \text{in to the } z\text{-direction with probability } p_{3\pm} + \sigma_3 q_3 \\
&= 0 && \text{with probability } -2\sigma q
\end{aligned} \tag{3}$$

with the constraint

$$p_{1+} + p_{1-} + p_{2+} + p_{2-} + p_{3+} + p_{3-} = 1. \tag{4}$$

Next if we choose step sizes $\delta x, \delta y, \delta z$ instead of ± 1 with the constraint

$$\lim \frac{(\delta x)^2}{\delta t} = D_1 \qquad \lim \frac{(\delta y)^2}{\delta t} = D_2 \qquad \lim \frac{(\delta z)^2}{\delta t} = D_3 \tag{5}$$

then we obtain

$$\begin{aligned}
\frac{\partial \pi}{\partial t} = \frac{1}{2} D_1 [2\sigma_1 q_1 + p_{1+} + p_{1-} - (p_{1+} - p_{1-})^2] \frac{\partial^2 \pi}{\partial x^2} \\
+ \frac{1}{2} D_2 [2\sigma_2 q_2 + p_{2+} + p_{2-} - (p_{2+} - p_{2-})^2] \frac{\partial^2 \pi}{\partial y^2} \\
+ \frac{1}{2} D_3 [2\sigma_3 q_3 + p_{3+} + p_{3-} - (p_{3+} - p_{3-})^2] \frac{\partial^2 \pi}{\partial z^2}
\end{aligned} \tag{6}$$

where it is tacitly assumed that π, which is a quaternion measure, is expressed in terms of σ matrices and the unit matrix. If π is post multiplied by an arbitrary spinor, we obtain a generalized version of the Pauli equation.

Instead of the random walk model, we can use as the starting point the Langevin equation

$$\mathrm{d}x^i = v^i \mathrm{d}\omega(t) + \tfrac{1}{2} v^j \delta_j v^i \mathrm{d}\omega^2(t) \tag{7}$$

where $\omega(t)$ is a Wiener process in the QMP framework and v^i is a function of x. If we let

$$E[\mathrm{d}x^i] = E[\Delta x^i] = D \tfrac{1}{2} V^j \delta_j V_i \tag{8}$$

$$E[\Delta x^i \Delta x^j] = D V^i V^j \tag{9}$$

then we arrive at the Fokker–Plank equation

$$\frac{\partial \pi(x,t)}{\partial t} = -\delta_i(\tfrac{1}{2}DV^j\delta_j V^i\pi) + \tfrac{1}{2}\delta_i\delta_j(DV^iV^j\pi). \tag{10}$$

If we choose

$$V^i = \sigma^i\beta(x) \qquad \text{(scalar)} \tag{11}$$

$$\text{or} \qquad V^i = \sigma^i\sigma^j\beta_j(x) \qquad \text{(vector)} \tag{12}$$

we obtain a generalized Pauli equation where π is a quaternion measure density and has the usual 2×2 matrix representation. Multiplication by an arbitrary spinor will lead to the familiar spinoral representation.

3. Dirac equation

Now we are well placed to deal with the process leading to the Dirac equation. Within the framework of QMP, we consider a Markov process $X(t)$ that satisfies the Langevin equation (see (2))

$$\mathrm{d}x^j = v^j \exp[\mathrm{i}\pi z(t)]\,\mathrm{d}t \tag{13}$$

where v_j as before is a function of x and $z(t)$ is a two-valued Markov process on $0, 1$ with transition rates λ_+ $(0\to 1)$ and λ_- $(1\to 0)$ per unit time. Thus the process $z(t)$ represents the transition from one to the other of the two helicity states. We assume $z(t)$ is independent of the process $x(t)$, the expectation values themselves being specified by

$$E[\mathrm{d}x^j|z(t)=0] = E[v^j(x)]\,\mathrm{d}t + o(\mathrm{d}t) \tag{14}$$

$$E[\mathrm{d}x^j|z(t)=1] = -E[v^j(x)]\,\mathrm{d}t + o(\mathrm{d}t). \tag{15}$$

Next we choose

$$E[v^j] = c\sigma^j \tag{16}$$

and adapt the Fokker–Plank method to yield

$$\frac{\partial \pi_+(x,t)}{\partial t} = -c\sigma\nabla\pi_+ - (\lambda_+\pi_+ - \lambda_-\pi_-) \tag{17}$$

where $\pi_+(x,t)\,\mathrm{d}x\,(\pi_-(x,t)\,\mathrm{d}x)$ represents quaternion measure that $x(t)$ lies in $(x, x+\mathrm{d}x)$ and $z(t)=0$ $(z(t)=1)$. In a similar way we obtain

$$\frac{\partial \pi_-(x,t)}{\partial t} = +c\sigma\nabla\pi_- - (\lambda_-\pi_- - \lambda_+\pi_+). \tag{18}$$

At this stage we postmultiply $\pi_\pm$ by an arbitrary 2-spinor $\chi_\pm$ to yield two-component objects (2-spinors) and we use the misnotation $\pi_\pm$ to denote the resulting 2-spinors.

If we now choose

$$\lambda_\pm = -\mathrm{i}\frac{mc^2}{h} \tag{19}$$

and set

$$\psi = \begin{pmatrix}\pi_+\\ \pi_-\end{pmatrix} \mathrm{e}^{-\mathrm{i}mc^2t/h} \tag{20}$$

we finally obtain the Dirac equation in 4-component form in the Weyl representation:

$$\mathrm{i}h\frac{\partial\psi}{\partial t} = mc^2\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\psi + \frac{hc}{\mathrm{i}}\begin{pmatrix}\hat{\sigma} & 0\\ 0 & -\hat{\sigma}\end{pmatrix}\nabla\psi. \tag{21}$$

It is worth noting that the plane wave solutions of the form

$$\psi = u \exp\left(-\frac{\mathrm{i}}{h}(Et - \hat{p}\hat{x})\right) \tag{22}$$

leads to

$$\begin{pmatrix} \pi_+ \\ \pi_- \end{pmatrix} = u\left(\exp -\frac{\mathrm{i}}{h}([E - mc^2] - \hat{p}\hat{x}\right) \tag{23}$$

a form which shows that the stationary state in the strict probabilistic sense is obtained if $E = mc^2$, in conformity with a similar result that can be obtained in the case of a free harmonic oscillator modelled in a complex measure theoretic framework. In this case the conditional measure density $f_2(x, \epsilon|x_0, t_0)$ satisfies

$$\frac{\partial f_2(x, t|x_0, t_0)}{\partial t} = \frac{\partial}{\partial x}(\mathrm{i}\omega x f_2) + \frac{\mathrm{i}h}{2m}\frac{\partial^2 f_2}{\partial x^2}. \tag{24}$$

The solution of f_2 discussed in [5] is connected to the ψ-function in a conventional wave mechanical treatment by

$$f_2(x, t|x_0 t_0)\mathrm{e}^{(m\omega/2h)(x^2 - x_0^2) - (\mathrm{i}\omega T/2)} = \psi(x, t|x_0, t_0) = \sum_{n=0}^{\infty} \mathrm{e}^{-\mathrm{i}E_n(T/h)}\phi_n(x)\phi_n^*(x_0) \tag{25}$$

where $T = t - t_0$. It is pertinent to note that $E_n = (n + \frac{1}{2})h\omega$ and this leads us to conclude that the energy in complex measure theory framework (CMTF) is now corresponding to the nth state and that the ground state is the only stationary state in the strict probabilistic sense. Moreover, we have

$$\lim_{T\to\infty} f_2 = \phi_0(x)\phi_0^*(x_0)\exp\left[-\frac{m\omega}{2h}(x^2 - x_0^2)\right] = \left(\frac{m\omega}{\pi h}\right)^{1/2}\exp\left(-\frac{m\omega x^2}{h}\right). \tag{26}$$

All other states ($n \neq 0$) are to be interpreted as quasistationary states with the use of the usual device $E_n \to E_n - \mathrm{i}\epsilon$. Now it is indeed possible to interpret the plane-wave-type solution from (21) analogously. Thus the only stationary state in the strict probabilistic sense is obtained when $E = mc^2$.

We next note that if the process $z(t)$ is replaced by a Poisson process $N(t)$, then

$$\mathrm{d}x^j = v^j \mathrm{e}^{\pi \mathrm{i} N(t)}\mathrm{d}t \tag{27}$$

provided we take the parameter λ of the Poisson process to be the same as λ_+ or λ_-; the original formulation is more general since λ_+ and λ_- can be distinct. Thus we can conclude that the Langevin equation description of the internal motion given by (11) or (20) together with the constraint that $z(t)$ is a Markov process (or $N(t)$ is a Poisson process) enable us to view the Dirac equation as a Markov process together with the stochastic flipping of the helicity. The internal motion with the σ matrix connection is essentially provided by the quaternion measure; the Poisson process $N(t)$ or the two-valued process $z(t)$ provides an easy interpretation of the change in the helicity due to internal motion. It is possible to introduce the potentials by the introduction of appropriate additional force terms in (20) or (11) as the case may be.

The choice of (11) or (20) rules out second/higher-order derivatives in as much as $E[(\Delta x)]$ is of order smaller than Δt, which is not the case in Langevin equations leading to Schrödinger/Pauli type of equations. Thus the Langevin equation (11) or (20) treats space and time on an equal footing, which is a mandatory requirement to obtain relativistic structures like the Dirac equation. Thus it is very likely that all spacetime symmetric physical structures have imbedded in them a universal Poisson process while non-relativistic structures have a universal Wiener process imbedded.

To introduce potentials we only need to modify (20) by

$$\mathrm{d}x^j = v^j \mathrm{e}^{\pi N(t)}\mathrm{d}t + \sum_1^\infty \frac{(\mathrm{d}N(t))^m}{m!}(g\hat{\nabla})^{m-1}g^j \tag{28}$$

where the special form involving powers of $\mathrm{d}N(t)$ is motivated by considerations of stochastic stability calculus (see [6, 7]). If we now proceed on lines exactly similar to the derivation of (15) and (16), we obtain

$$\frac{\partial\pi_\pm(x,t)}{\partial t} = \mp c(\sigma\nabla)\pi_\pm - \lambda(\pi_\pm - \pi_\mp + \lambda(\nabla\pi_\pm)\hat{G} \tag{29}$$

where

$$G^n = \sum_1 \frac{1}{m!}(g\nabla)^{m-1}g^i. \tag{30}$$

With $\pi_\pm$ reinterpreted as before, and on substitution of (18), we obtain

$$\mathrm{i}h\frac{\partial\psi}{\partial t} = mc^2\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\psi + c\begin{pmatrix}\hat{\sigma} & 0\\ 0 & -\hat{\sigma}\end{pmatrix}\frac{h}{c}\nabla\psi + \mathrm{i}h\nabla\psi\cdot\boldsymbol{G}. \tag{31}$$

Using the identity

$$\nabla\psi\cdot\boldsymbol{G} = -\psi\nabla\cdot\boldsymbol{G} + \nabla(\psi\boldsymbol{G}) \tag{32}$$

and integrating over all space, we observe that the term does not contribute provided proper boundary conditions are applied. Thus (31) can be regarded as a general form of Dirac equation with potentials.

4. Summary and conclusion

Since diffusion in the classical context is a non-relativistic way of describing motions, we have resorted to a spacetime symmetric mechanism. This simple modification described above leads to the Dirac equation in its simplest form. On the other hand, a quaternionic version of a generalized random walk leads to a generalized Pauli system.

It is pertinent to point out that an attempt to derive the Dirac equation was made by Gaveau *et al* [8]. Their main concern was to relate the Dirac equation to the Telegrapher's equation by an analytic continuation. However, our considerations establish that we have to go beyond the framework of complex measures to provide a satisfactory derivation of the Dirac equation from a probabilistic view point.

The use of 2×2 matrices $\mathrm{i}\sigma$ together with unity, which form the algebra of quaternions, may suggest some relationship to the generalization of quantum mechanics over the field of quaternions [9]. However, the matrices here are Dirac algebra generators and they are connected with spin and Lorentz transformations. Therefore the similarity is only superficial.

References

[1] Srinivasan S K and Sudarshan E C G 1994 *J. Phys. A: Math. Gen.* **27** 517
[2] Ash R B 1972 *Real Analysis and Probability* (New York: Academic)
[3] Smith D E 1959 *A Source Book of Mathematics* vol 2 (New York: McGraw Hill) p 677
[4] Pitt H R *Integration, Measure and Probability* (London: Oliver and Boyd)
[5] Srinivasan S K 1995 *Topics in Mathematical Physics – Professor R Vasudevan Memorial Volume 1* ed Rao K Srinivasa (Singapore: World Scientific) to be published

[6] Mcshane E J 1974 *Stochastic Calculus and Models* (New York: Academic)
[7] Srinivasan S K 1982 *Stochastic Calculus and Models and Their Applications to Some Problems of Physics (Lecture Notes Series 15)* (Singapore: National University of Singapore)
[8] Gaveau B, Jacobson T, Kac M and Schulman L S 1984 *Phys. Rev. Lett.* **53** 419
[9] Adler S L 1986 *Phys. Rev.* D **34** 1871; 1986 *Comm. Math. Phys.* **104** 611; 1988 *Phys. Rev.* D **37** 3654
Horwitz L P 1993 *J. Math. Phys.* **34** 3405