Eur. J. Phys. 10 (1989) 188–193. Printed in the UK

A quaternion expression for the quantum mechanical probability and current densities

W Gough

Department of Physics, University College, Cardiff, UK

Received 30 August 1988

Abstract. A simple quaternion expression is proposed for the quantum mechanical four-current of a single particle. It is shown that the equation of continuity is obeyed and the behaviour under a Lorentz transformation is correct. The expression is almost equivalent to that emanating from the Dirac theory. It differs, however, from those of other authors who have considered the quaternion formulation of relativistic quantum mechanics.

Résumé. Une expression quaternionique est proposée pour le quadri-courant d'une particule. On montre qu'elle satisfait l'équation de continuité et que son comportement sous une transformation de Lorentz est correct. L'expression est presque équivalente à celle obtenue à partir de la théorie de Dirac. Elle diffère cependant de celles d'autres auteurs qui ont considéré la formulation quaternionique de la mécanique quantique relativiste.

1. Introduction

The interpretation of the wavefunction Ψ for a single particle in non-relativistic quantum mechanics is that

$$\rho = \Psi^* \Psi \tag{1}$$

represents the probability density. A straightforward analysis to be found in many textbooks (e.g. Schiff 1955 p.24) reveals that

$$\operatorname{div} \boldsymbol{j} = -\partial \rho / \partial t \tag{2}$$

where

$$\boldsymbol{j} = \frac{i\hbar}{2m} (\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi) \tag{3}$$

is interpreted as the probability current density which satisfies the equation of continuity (2). This latter is of course familiar in the context of electrical theory, where ρ and $\boldsymbol{j}$ represent the charge and current densities, the expresses the condition for conservation of electric charge. Under a Lorentz transformation (hereafter denoted LT) ρ and $\boldsymbol{j}$ follow the same equations of transformation as those for t and $\boldsymbol{r}(=(x, y, z))$; that is to say, $(\boldsymbol{j}, ic\rho)$ is a four-vector (e.g. Lorrain and Corson 1970).

In this paper we shall use the quaternion formulation of the relativistic wave equation for an electron (or positron) to show that there is a very simple expression for the four-vector $(\boldsymbol{j}, ic\rho)$ which obeys the equation of continuity (2) and has the correct LT. Surprisingly, it appears to be original and to have escaped the attention of other authors who have sought quaternion expressions for ρ and $\boldsymbol{j}$.

2. The relativistic wave equation for the electron

In the Dirac theory (Berestetskii *et al* 1971) the wavefunction Ψ has four components and as such is expressible as a column vector $(\xi^1, \xi^2, \eta_1, \eta_2)$, here written as a row vector for economy of space. The

0143-0807/89/030188+06 \$02.50

version of the Dirac equation known as the spinor (or symmetrical) representation may be written

$$\mathrm{i}\hbar\left(\frac{\partial}{\partial t}+c\boldsymbol{\sigma}\cdot\nabla\right)\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix}=mc^2\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}$$
$$\mathrm{i}\hbar\left(\frac{\partial}{\partial t}-c\boldsymbol{\sigma}\cdot\nabla\right)\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=mc^2\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix} \tag{4}$$

where

$$\sigma_x=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\qquad \sigma_y=\begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}\qquad \sigma_z=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

are the Pauli spin matrices.

The equation of continuity (2) is satisfied if ρ is still given by (1) but now $\boldsymbol{j}=c\Psi^*\boldsymbol{\alpha}\Psi$ where, in the symmetrical representation,

$$\boldsymbol{\alpha}=\begin{pmatrix}\boldsymbol{\sigma} & 0\\ 0 & -\boldsymbol{\sigma}\end{pmatrix}$$

its components therefore being 4×4 matrices.

Here, however, we shall use the quaternion approach to the relativistic theory of the electron/positron. This offers some advantages over the Dirac formulation, namely: (i) It is unnecessary to use matrices with their attendant conceptual difficulties, for example, that of 'spin space'. Instead, specific expressions are used for spinors which make them immediately appreciable. In the opinion of the author, the quaternion formulation is in some ways aesthetically more satisfying than the Dirac formulation. (ii) The quaternion approach is easier to understand, making it particularly appealing both for teacher and learner. Indeed, quaternions are no more difficult than vectors.

The mathematics of quaternions has been expounded in a series of papers (Gough 1984, 1986, 1987) with particular reference to quantum mechanics. The reader is referred to these for further details. For convenience we repeat and develop the relevant points in Appendix 1, where the terminology is also defined.

The Dirac equation for a free electron (or positron) of rest mass m is expressible as two simultaneous equations (e.g. Brown 1962) which have the quaternion form (Gough 1987, with notation slightly altered to accord with Berestetskii *et al* 1971)

$$\left(\frac{\mathrm{i}\hbar}{c}\frac{\partial}{\partial t}-\hbar\nabla_{\mathrm{q}}\right)\xi=mc\eta$$
$$\left(\frac{\mathrm{i}\hbar}{c}\frac{\partial}{\partial t}+\hbar\nabla_{\mathrm{q}}\right)\eta=mc\xi. \tag{5}$$

The wavefunction can be regarded as having two components ξ and η. More generally, in the presence of an external field with scalar potential V and vector potential (A_x, A_y, A_z),

$$\left(\frac{\mathrm{i}\hbar}{c}\frac{\partial}{\partial t}-\hbar\nabla_{\mathrm{q}}-\frac{QV}{c}+\mathrm{i}QA_{\mathrm{q}}\right)\xi=mc\eta$$
$$\left(\frac{\mathrm{i}\hbar}{c}\frac{\partial}{\partial t}+\hbar\nabla_{\mathrm{q}}-\frac{QV}{c}-\mathrm{i}QA_{\mathrm{q}}\right)\eta=mc\xi \tag{6}$$

where $A_{\mathrm{q}}=A_x\imath+A_y\jmath+A_z\kappa$ and Q is the charge $-e$ (electron) or $+e$ (positron).

3. Four-vector for probability and probability current densities

The quaternionic form of the four-current $(\boldsymbol{j}, \mathrm{i}c\rho)$ can be defined as $c\rho\pm\mathrm{i}j_q$, where $j_{\mathrm{q}}=j_x\imath+j_y\jmath+j_z\kappa$ (cf. Gough 1987 for example). From equation (A8)

$$\left(\nabla_{\mathrm{q}}-\frac{\mathrm{i}}{c}\frac{\partial}{\partial t}\right)(c\rho+\mathrm{i}j_{\mathrm{q}})$$
$$=c\,\mathrm{grad}\,\rho-\frac{\mathrm{i}\partial\rho}{\partial t}+\mathrm{i}(-\mathrm{div}\,j_{\mathrm{q}}+\mathrm{curl}\,j_{\mathrm{q}})+\frac{1}{c}\frac{\partial j_{\mathrm{q}}}{\partial t}.$$

The scalar part of the RHS is zero from the equation of continuity (2), which can therefore be written in the form

$$S\left[\left(\nabla_{\mathrm{q}}-\frac{\mathrm{i}}{c}\frac{\partial}{\partial t}\right)(c\rho+\mathrm{i}j_{\mathrm{q}})\right]=0 \tag{7}$$

We now propose that ρ and j_{q} be expressible by the following simple expression involving the suitably normalised wavefunctions ξ and η

$$\rho+\mathrm{i}\frac{j_{\mathrm{q}}}{c}=\xi\xi^*+\eta^*\bar{\eta}. \tag{8}$$

The RHS is its own full conjugate and as such is indeed expressible as a real scalar plus a pure imaginary vector (Appendix 1).

For it to be acceptable, a necessary condition is that the equation of continuity (7) be satisfied, i.e.

$$S\left[\left(\nabla_q - \frac{i}{c}\frac{\partial}{\partial t}\right)(\xi\bar{\xi}^* + \eta^*\bar{\eta})\right] = 0 \qquad (9)$$

which we now prove. From (A7), the LHS is

$$S\left[(\nabla_q\xi)\bar{\xi}^* - \xi(\overline{\nabla_q\xi^*}) - \frac{i}{c}\left(\frac{\partial\xi}{\partial t}\bar{\xi}^* + \xi\frac{\partial\bar{\xi}^*}{\partial t}\right)\right.$$

$$\left. + (\nabla_q\eta^*)\bar{\eta} - \eta^*(\overline{\nabla_q\eta}) - \frac{i}{c}\left(\frac{\partial\eta^*}{\partial t}\bar{\eta} + \eta^*\frac{\partial\bar{\eta}}{\partial t}\right)\right]$$

$$= \hbar^{-1}S\left[\left(\frac{-QV\xi}{c} + iQA_q\xi - mc\eta\right)\bar{\xi}^*\right.$$

$$+ \xi\left(\frac{QV}{c}\bar{\xi}^* - iQ\bar{\xi}^*A_q + mc\bar{\eta}^*\right)$$

$$+ \left(\frac{QV}{c}\eta^* - iQA_q\eta^* + mc\xi^*\right)\bar{\eta}$$

$$\left. - \eta^*\left(\frac{QV}{c}\bar{\eta} - iQ\bar{\eta}A_q + mc\bar{\xi}\right)\right].$$

We have used here equations (6), their conjugates, (A5) and $A_q = -\bar{A}_q$. After some cancellation, we find that the LHS is equal to

$$\hbar^{-1}S[iQ(A_q\xi\bar{\xi}^* - \xi\bar{\xi}^*A_q - A_q\eta^*\bar{\eta} + \eta^*\bar{\eta}A_q) + mc(\xi\bar{\eta}^* + \xi^*\bar{\eta} - \eta\bar{\xi}^* - \eta^*\bar{\xi})].$$

But from (A4), $S(A_q\xi\bar{\xi}^* - \xi\bar{\xi}^*A_q)$ and $S(-A_q\eta^*\bar{\eta} + \eta^*\bar{\eta}A_q)$ are both zero. In addition, the second term in parentheses can be written from (A5) in the form $(\xi\bar{\eta}^* + \xi^*\bar{\eta})$ − quat. conj., and is therefore a pure vector. It follows then that the LHS of (9) is indeed zero, as required.

4. Relativistic transformation

In Appendix 2 we give the necessary background to deduce how the various terms in (6) behave under a LT. From (A14) and (A15) the transformed equations read

$$\exp(-\tfrac{1}{2}i_r\theta)\left(\frac{i\hbar}{c}\frac{\partial}{\partial t'} - \hbar\nabla'_q - \frac{QV'}{c} + iQA'_q\right) \times \exp(-\tfrac{1}{2}i_r\theta)\xi = mc\eta$$

$$\exp(\tfrac{1}{2}i_r\theta)\left(\frac{i\hbar}{c}\frac{\partial}{\partial t'} + \hbar\nabla'_q - \frac{QV'}{c} - iQA'_q\right) \times \exp(\tfrac{1}{2}i_r\theta)\eta = mc\xi.$$

Pre-multiplication by $\exp(\frac{1}{2}i_r\theta)$ and $\exp(-\frac{1}{2}i_r\theta)$, respectively, and comparison with (6) reveals that the LTs of ξ and η are

$$\xi' = \exp(-\tfrac{1}{2}i_r\theta)\xi \qquad \eta' = \exp(\tfrac{1}{2}i_r\theta)\eta.$$

We can now deduce the LT of our proposed expression (8) for the probability density four-current (divided by c) using equations (A5), (A6)

$$\rho' + \frac{ij'_q}{c} = \xi'\bar{\xi}'^* + \eta'^*\bar{\eta}'$$

$$= \exp(-\tfrac{1}{2}i_r\theta)\xi\bar{\xi}^*\exp(-\tfrac{1}{2}i_r\theta) + \exp(-\tfrac{1}{2}i_r\theta)\eta^*\bar{\eta}\exp(-\tfrac{1}{2}i_r\theta)$$

$$= \exp(-\tfrac{1}{2}i_r\theta)\left(\rho + \frac{ij_q}{c}\right)\exp(-\tfrac{1}{2}i_r\theta).$$

Comparison with (A12) shows that the transformation is indeed appropriate to a four-current, adding further weight to the validity of our proposal.

5. Probability and current densities in terms of φ and χ. Non-relativistic approximation

The wave equations (6) may alternatively be written by defining the functions

$$\varphi = \frac{1}{\sqrt{2}}(\xi + \eta) \qquad \chi = \frac{1}{\sqrt{2}}(\xi - \eta)$$

whence

$$(\hbar\nabla_q - iQA_q)\varphi = \left(\frac{E - QV}{c} + mc\right)\chi \qquad (10a)$$

$$(\hbar\nabla_q - iQA_q)\chi = \left(\frac{E - QV}{c} - mc\right)\varphi \qquad (10b)$$

since $E = i\hbar\partial/\partial t$. φ and χ are the quaternion form of the components of the Dirac wavefunction Ψ in the standard representation.

In terms of φ and χ, the proposed equation (8) becomes

$$\rho + ij_q/c = \xi\bar{\xi}^* + \eta^*\bar{\eta}$$

$$= \tfrac{1}{2}[(\varphi + \chi)(\bar{\varphi}^* + \bar{\chi}^*) + (\varphi^* - \chi^*)(\bar{\varphi} - \bar{\chi})]$$

$$= \tfrac{1}{2}[(\varphi\bar{\varphi}^* + \chi\bar{\chi}^* + \text{complex conj.}) + (\varphi\bar{\chi}^* + \chi\bar{\varphi}^* - \text{quat. conj.})].$$

The first bracket is a real scalar and the second a pure vector, whence

$$\begin{aligned} \rho &= \tfrac{1}{2}(\varphi\bar{\varphi}^* + \varphi^*\bar{\varphi} + \chi\bar{\chi}^* + \chi^*\bar{\chi}) \\ j_q &= \tfrac{1}{2}ic(-\varphi\bar{\chi}^* + \varphi^*\bar{\chi} - \chi\bar{\varphi}^* + \chi^*\bar{\varphi}). \end{aligned} \qquad (11)$$

In the non-relativistic approximation $E/c - mc$ is very small and hence from (10a) $\chi \ll \varphi$. φ is the conventional wavefunction in non-relativistic theory, which is of course always taken to be a scalar quantity. If we assume $A_q = 0$ for simplicity, then (10a) and (A8) lead to grad $\varphi \approx (2mc/\hbar)\chi$. Therefore

χ is a pure vector and so $\bar{\chi} = -\chi$. Equations (11) become $\rho \approx \varphi^* \varphi$, which is manifestly correct, and

$$j = \mathrm{i}c(\varphi\chi^* - \varphi^*\chi) = \frac{\mathrm{i}\hbar}{2m}(\varphi \operatorname{grad} \varphi^* - \varphi^* \operatorname{grad} \varphi).$$

This is in accord with the conventional expression (3).

6. Comparison with other expressions for ρ and j_{q}

We have shown above that in the non-relativistic limit expression (8) reduces to the conventional equations (1) and (3) for ρ and $\boldsymbol{j}$. Comparison may also be made with the Dirac relativistic theory (with $V = \boldsymbol{A} = 0$ for simplicity). According to this, the four-current is expressible as

$$j^\mu = \bar{\Psi}\gamma^\mu\Psi \tag{12}$$

where $j^\mu = (\rho, j_x/c, j_y/c, j_z/c)$, $\gamma^\mu = (\gamma^0, \gamma_x, \gamma_y, \gamma_z)$ and $\bar{\Psi}(=\Psi^*\gamma^0)$ is the Dirac conjugate of Ψ. In the symmetrical representation

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \gamma^0\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix}.$$

The quaternion spinors are the eigenfunctions of the spin operator $S_z/\hbar = \frac{1}{2}\mathrm{i}\boldsymbol{k}$, namely $1 + \mathrm{i}\boldsymbol{k}$ or $\mathrm{i}\boldsymbol{i} - \boldsymbol{j}$ ($m_s = \frac{1}{2}$, spin up), and $\mathrm{i}\boldsymbol{i} + \boldsymbol{j}$ or $1 - \mathrm{i}\boldsymbol{k}$ ($m_s = -\frac{1}{2}$, spin down).

We expand ξ and η in terms of these spinors as follows

$$\xi = \frac{1}{\sqrt{2}}[\xi^1(1 + \mathrm{i}\boldsymbol{k}) + \xi^2(\mathrm{i}\boldsymbol{i} + \boldsymbol{j}) + \xi'^1(\mathrm{i}\boldsymbol{i} - \boldsymbol{j}) + \xi'^2(1 - \mathrm{i}\boldsymbol{k})]$$

$$\eta = \frac{1}{\sqrt{2}}[\eta_1(1 + \mathrm{i}\boldsymbol{k}) + \eta_2(\mathrm{i}\boldsymbol{i} + \boldsymbol{j}) + \eta_1'(\mathrm{i}\boldsymbol{i} - \boldsymbol{j}) + \eta_2'(1 - \mathrm{i}\boldsymbol{k})] \tag{13}$$

where $\xi^1, \ldots, \eta_2$ do not contain $\boldsymbol{i}, \boldsymbol{j}$ or $\boldsymbol{k}$. Substitution into (5) gives differential equations for ξ^1, ξ^2, η_1 and η_2 which are identical to the expanded Dirac equation (4), and so do the equations for the primed quantities. Equations (13) therefore represent a linear sum of two independent Dirac-type wavefunctions. There are eight components of the quaternion wavefunction, whereas the Dirac wavefunction involves only four.

Whether a particle requires eight components to specify it is doubtful. The success of the Dirac theory would indicate that four components are sufficient, in which case we could ignore the last two terms in each of (13). This conclusion appears to be supported by an experiment to look for 'quaternion phase shifts' (Kaiser *et al* 1984).

We have considered each of the solutions separately. A full analysis (not given here) reveals that the equation (12) for the four-current is in agreement with the Dirac theory. However, our proposed equation (8) for the four-current is also appropriate to a superposition of these solutions and is therefore more general than the Dirac theory expression. We now consider other authors who have used the quaternion formulation of the relativistic wave equation.

Conway's approach (1937, 1948) has the advantage of describing the particle by a single wave equation. That given in the 1937 reference may be written in our nomenclature

$$(\hbar\nabla_{\mathrm{q}} - \mathrm{i}QA_{\mathrm{q}})\Psi\boldsymbol{i} - \left(\mathrm{i}\hbar\frac{\partial}{\partial t} - \frac{QV}{c}\right)\Psi\mathrm{i}\boldsymbol{j} + mc\Psi\boldsymbol{k} = 0.$$

It does, however, suffer from the disadvantage of introducing the unit elements $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ directly, thereby destroying the symmetry which is present in (6). As is readily shown, one possible substitution by which these latter reduce to Conway's equation is

$$\Psi = -\xi(\mathrm{i}\boldsymbol{i} + \boldsymbol{j}) + \eta(1 - \mathrm{i}\boldsymbol{k}).$$

Conway also gives (essentially) $\rho = S(\Psi\bar{\Psi}^*)$, $j_{\mathrm{q}} = V(\Psi\boldsymbol{i}\bar{\Psi}^*)$, which are, therefore, in our notation

$$\begin{aligned}
\rho &= 2S\{\xi(1 - \mathrm{i}\boldsymbol{k})\bar{\xi}^* + \eta(1 - \mathrm{i}\boldsymbol{k})\bar{\eta}^*\} \\
&= \xi(1 - \mathrm{i}\boldsymbol{k})\bar{\xi}^* + \xi^*(1 + \mathrm{i}\boldsymbol{k})\bar{\xi} \\
&\quad + \eta(1 - \mathrm{i}\boldsymbol{k})\bar{\eta}^* + \eta^*(1 + \mathrm{i}\boldsymbol{k})\bar{\eta} \\
-\mathrm{i}j_{\mathrm{q}} &= 2V\{\xi(1 - \mathrm{i}\boldsymbol{k})\bar{\eta}^* + \eta(1 - \mathrm{i}\boldsymbol{k})\bar{\xi}^*\} \\
&= \xi(1 - \mathrm{i}\boldsymbol{k})\bar{\eta}^* - \xi^*(1 + \mathrm{i}\boldsymbol{k})\bar{\eta} \\
&\quad + \eta(1 - \mathrm{i}\boldsymbol{k})\bar{\xi}^* - \eta^*(1 + \mathrm{i}\boldsymbol{k})\bar{\xi}
\end{aligned}$$

which are quite different from (8).

Edmonds (1972) adopts the formulism frequently used in quaternion theory in which e_0, e_1, e_2 and e_3 are isomorphic with 1, $\mathrm{i}\boldsymbol{i}$, $\mathrm{i}\boldsymbol{j}$ and $\mathrm{i}\boldsymbol{k}$ respectively. His form of the wave equations is isomorphic with ours, with his ψ_{v} and ψ_{a} equivalent to our ξ and η. His final expressions for ρ and j_{q} are in our nomenclature

$$\rho = \tfrac{1}{2}(\xi\bar{\xi}^* + \bar{\xi}^*\xi + \bar{\eta}\eta^* + \bar{\eta}^*\eta)$$

$$j_{\mathrm{q}} = \tfrac{1}{2}c(\bar{\xi}\boldsymbol{i}\xi^* - \bar{\xi}^*\boldsymbol{i}\xi + \bar{\eta}^*\boldsymbol{i}\eta - \bar{\eta}\boldsymbol{i}\eta^*)\boldsymbol{i} + \ldots$$

which bear only a little resemblance to (8).

In a later paper, Edmonds (1978) gives a more elaborate expression for the current four-vector in terms of the Dirac matrices, which appear to resemble the expressions given by Conway.

7. Conclusion

We have produced a simple expression, believed to be original, for the probability four-current in relativistic quantum theory, which appears to be satisfactory in all respects. Our expression is not equivalent to others derived by quaternion theory which have appeared in the literature, but the latter are much less elegant than that proposed here. It is (almost) equivalent to the Dirac four-current, but in the opinion of the author, ours is more immediately appealing than the latter. We contend that the use of quaternions in relativity and quantum mechanics can be powerful and elegant, often more aesthetically pleasing and comprehensible than conventional approaches.

Appendix 1. The elements of quaternions

In quaternion theory, there are three anticommuting 'values' of $\sqrt{-1}$, namely $\imath$, $\jmath$ and κ defined by

$$\imath^2=\jmath^2=\kappa^2=-1 \qquad \imath\jmath=-\jmath\imath=\kappa$$
$$\jmath\kappa=-\kappa\jmath=\imath \qquad \kappa\imath=-\imath\kappa=\jmath.$$

Complex quaternion theory also introduces the 'usual value' i for $\sqrt{-1}$, which commutes with all quantities and causes no complications. A quaternion may be expressed as

$$q=a_0+a_1\imath+a_2\jmath+a_3\kappa \tag{A1}$$

where $a_0 \ldots a_3$ are ordinary complex numbers in the conventional sense, expressible as $\alpha_0+\mathrm{i}\beta_0$ (α_0 and β_0 real) etc.

Traditionally, a_0 and $a_1\imath+a_2\jmath+a_3\kappa$ are called the scalar and vector parts of q, denoted by $S(q)$ and $V(q)$ respectively.

We shall use the following terminology and definitions of conjugates:

(i) The complex conjugate of q is

$$q^*=a_0^*+a_1^*\imath+a_2^*\jmath+a_3^*\kappa$$

where $a_0^*=\alpha_0-\mathrm{i}\beta_0$ etc. Under this operation, $\mathrm{i}\to-\mathrm{i}$, $\imath\to\imath$, $\jmath\to\jmath$, $\kappa\to\kappa$.

(ii) The quaternion conjugate of q is

$$\bar{q}=a_0-a_1\imath-a_2\jmath-a_3\kappa$$

Here $\mathrm{i}\to\mathrm{i}$, $\imath\to-\imath$, $\jmath\to-\jmath$, $\kappa\to-\kappa$.

(iii) That formed by combining these operations we shall call, in the absence of standard terminology, the full conjugate

$$\bar{q}^*=a_0^*-a_1^*\imath-a_2^*\jmath-a_3^*\kappa \tag{A2}$$

that is i, $\imath$, $\jmath$ and κ all change sign.

The product of two quaternions $q_1=a_0+a_1\imath+a_2\jmath+a_3\kappa$ and $q_2=b_0+b_1\imath+b_2\jmath+b_3\kappa$ is

$$q_1q_2=a_0b_0-a_1b_1-a_2b_2-a_3b_3$$
$$+\imath(a_0b_1+a_1b_0+a_2b_3-a_3b_2)+\ldots \tag{A3}$$

from which the following theorems are easily derived:

$$S(q_1q_2)=S(q_2q_1) \tag{A4}$$
$$\overline{q_1q_2}=\bar{q}_2\bar{q}_1 \tag{A5}$$
$$(\overline{q_1q_2})^*=\bar{q}_2^*\bar{q}_1^*. \tag{A6}$$

Of particular interest is the product $q\bar{q}^*$. From (A1) and (A2) we see that the *scalar part is real* and the *vector part is pure imaginary*. This conclusion also follows from (A6) which shows that $q\bar{q}^*$ is its own full conjugate.

We shall also need to evaluate the result of $\nabla_q(=\imath\,\partial/\partial x+\jmath\,\partial/\partial y+\kappa\,\partial/\partial z)$ acting on a quaternion product. Because of the non-commutative properties of quaternions, it would be incorrect to write $\nabla_q(q_1q_2)=q_1(\nabla_q q_2)+q_2(\nabla_q q_1)$. The correct expression is rather cumbersome. However, we shall be concerned only with the scalar part, which we now deduce. From (A3)

$$S[\nabla_q(q_1q_2)]=-\frac{\partial}{\partial x}(a_0b_1+a_1b_0+a_2b_3-a_3b_2)$$
$$-\frac{\partial}{\partial y}(\ldots)-\frac{\partial}{\partial z}(\ldots).$$

But

$$S[(\nabla_q q_1)q_2]$$
$$=-\left(\frac{\partial a_0}{\partial x}b_1+\frac{\partial a_1}{\partial x}b_0+\frac{\partial a_2}{\partial x}b_3-\frac{\partial a_3}{\partial x}b_2\right)-\ldots$$

and therefore

$$S[-\nabla_q(q_1q_2)+(\nabla_q q_1)q_2]$$
$$=\left(a_0\frac{\partial b_1}{\partial x}+a_1\frac{\partial b_0}{\partial x}+a_2\frac{\partial b_3}{\partial x}-a_3\frac{\partial b_2}{\partial x}\right)+\ldots$$
$$=S[(a_0+a_1\imath+a_2\jmath+a_3\kappa)$$
$$\left(\frac{\partial b_1}{\partial x}-\imath\frac{\partial b_0}{\partial x}-\jmath\frac{\partial b_3}{\partial x}+\kappa\frac{\partial b_2}{\partial x}\right)+\ldots$$
$$=S\left[q_1\overline{\left(\imath\frac{\partial}{\partial x}+\jmath\frac{\partial}{\partial y}+\kappa\frac{\partial}{\partial z}\right)(b_0-b_1\imath-b_2\jmath-b_3\kappa)}\right].$$

It follows that

$$S[\nabla_q(q_1q_2)]=S[(\nabla_q q_1)q_2-q_1(\overline{\nabla_q\bar{q}_2})].$$

In particular,

$$S[\nabla_q(q\bar{q}^*)]=S[(\nabla_q q)\bar{q}^*-q(\overline{\nabla_q q^*})]. \tag{A7}$$

Finally, we note that the result of ∇_q operating on a quaternion can be elegantly expressed in vector operator form, since

$$\nabla_q q=\left(\frac{\imath\,\partial}{\partial x}+\frac{\jmath\,\partial}{\partial y}+\frac{\kappa\,\partial}{\partial z}\right)(a_0+a_1\imath+a_2\jmath+a_3\kappa)$$
$$=\nabla_q a_0-\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)+\imath\left(\frac{\partial a_3}{\partial y}-\frac{\partial a_2}{\partial z}\right)+\ldots$$

which can be written, with obvious notation,

$$\nabla_q q=\text{grad (scalar part)}-\text{div (vector part)}$$
$$+\text{curl (vector part)}. \tag{A8}$$

Appendix 2. Relativistic transformations by quaternions

It has long been recognised that quaternion algebra is well suited to the formulation of special relativity (Silberstein 1912, 1913, Dirac 1945, Rastall 1964, Synge 1972). The displacement–time four-vector may be expressed as the single entity $ct \pm \mathrm{i}(\boldsymbol{\imath}x + \boldsymbol{\jmath}y + \boldsymbol{k}z)$, with corresponding expressions for other physical four-vectors. We have adopted a convention in which the scalar part is real and the vector part is pure imaginary. (This convention is in some ways preferable to the alternative notion of a 'minquat' (Synge 1972), in which the scalar part is pure imaginary and the vector part is real.)

Consider the Lorentz transformations (LTS) $ct' = \gamma(ct - \beta x)$, $x' = \gamma(x - \beta ct)$, $y' = y$, $z' = z$, where $\beta = v/c$, $\gamma = (1-\beta^2)^{-1/2}$. The first two equations may be written $ct' = ct\cosh\theta - x\sinh\theta$, $x' = x\cosh\theta - ct\sinh\theta$ where $\cosh\theta = \gamma$, $\sinh\theta = \beta\gamma$ (so $\cosh^2\theta - \sinh^2\theta = 1$ as required). Therefore

$$
\begin{aligned}
ct' \pm \mathrm{i}\boldsymbol{\imath}x' &= (ct \pm \mathrm{i}\boldsymbol{\imath}x)(\cosh\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\theta) \\
&= (\cosh\tfrac{1}{2}\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta) \\
&\quad \times (ct \pm \mathrm{i}\boldsymbol{\imath}x)(\cosh\tfrac{1}{2}\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta) \qquad \text{(A9)}
\end{aligned}
$$

since $\cosh\theta = \cosh^2\tfrac{1}{2}\theta + \sinh^2\tfrac{1}{2}\theta$, $\sinh\theta = 2\sinh\tfrac{1}{2}\theta\cosh\tfrac{1}{2}\theta$ and all terms in the expansion commute. We note too that

$$
\begin{aligned}
&(\cosh\tfrac{1}{2}\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta)(\mathrm{i}\boldsymbol{\jmath}y + \mathrm{i}\boldsymbol{k}z) \\
&\qquad \times (\cosh\tfrac{1}{2}\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta) \\
&= (\cosh\tfrac{1}{2}\theta \mp \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta)(\cosh\tfrac{1}{2}\theta \pm \mathrm{i}\boldsymbol{\imath}\sinh\tfrac{1}{2}\theta) \\
&\qquad \times (\mathrm{i}\boldsymbol{\jmath}y + \mathrm{i}\boldsymbol{k}z) = \mathrm{i}\boldsymbol{\jmath}y + \mathrm{i}\boldsymbol{k}z = \mathrm{i}\boldsymbol{\jmath}y' + \mathrm{i}\boldsymbol{k}z'. \qquad \text{(A10)}
\end{aligned}
$$

Now the exponential of a quaternion can be defined via the usual series expansion, whence

$$
\begin{aligned}
\exp(\mathrm{i}\boldsymbol{\imath}\theta) &= 1 + \mathrm{i}\boldsymbol{\imath}\theta + \frac{(\mathrm{i}\boldsymbol{\imath}\theta)^2}{2!} + \ldots \\
&= 1 + \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \ldots + \mathrm{i}\boldsymbol{\imath}\left(\theta + \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \ldots\right) \\
&= \cosh\theta + \mathrm{i}\boldsymbol{\imath}\sinh\theta.
\end{aligned}
$$

Using (A9), (A10) and putting $r'_{\mathrm{q}} = \boldsymbol{\imath}x' + \boldsymbol{\jmath}y' + \boldsymbol{k}z'$, $r_{\mathrm{q}} = \boldsymbol{\imath}x + \boldsymbol{\jmath}y + \boldsymbol{k}z$, we have, as given by Silberstein (1912),

$$
ct' \pm \mathrm{i}r'_{\mathrm{q}} = \exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)(ct \pm \mathrm{i}r_{\mathrm{q}})\exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)
$$

which is a very elegant expression for the LT. Of course it will also pertain *mutatis mutandis* to all physical four-vectors, particularly those for the electromagnetic four-potential $(\boldsymbol{A}, \mathrm{i}V/c)$ and the four-current density $(\boldsymbol{j}, \mathrm{i}c\rho)$,

$$
V'/c \pm \mathrm{i}A'_{\mathrm{q}} = \exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)(V/c \pm \mathrm{i}A_{\mathrm{q}})\exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta) \qquad \text{(A11)}
$$

$$
c\rho' \pm \mathrm{i}j'_{\mathrm{q}} = \exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)(c\rho \pm \mathrm{i}j_{\mathrm{q}})\exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta). \qquad \text{(A12)}
$$

Likewise, the four-vector equivalent of ∇_{q} transforms as follows

$$
\frac{\mathrm{i}}{c}\frac{\partial}{\partial t'} \pm \nabla'_{\mathrm{q}} = \exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)\left(\frac{\mathrm{i}}{c}\frac{\partial}{\partial t} \pm \nabla_{\mathrm{q}}\right)\exp(\mp\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta). \qquad \text{(A13)}
$$

The inverse transformations of (A11) and (A13), needed in §4, are

$$
V/c \pm \mathrm{i}A_{\mathrm{q}} = \exp(\pm\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)(V'/c \pm \mathrm{i}A'_{\mathrm{q}})\exp(\pm\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta) \qquad \text{(A14)}
$$

$$
\frac{\mathrm{i}}{c}\frac{\partial}{\partial t} \pm \nabla_{\mathrm{q}} = \exp(\pm\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta)\left(\frac{\mathrm{i}}{c}\frac{\partial}{\partial t'} \pm \nabla'_{\mathrm{q}}\right)\exp(\pm\tfrac{1}{2}\mathrm{i}\boldsymbol{\imath}\theta). \qquad \text{(A15)}
$$

References

Berestetskii V B, Lifshitz E M and Pitaevskii L P 1971 *Relativistic Quantum Theory* in *Course of Theoretical Physics* ed. L D Landau and E M Lifshitz vol. 4 part 1 (Oxford: Pergamon)

Brown L M 1962 *Lectures in Theoretical Physics* (Boulder) vol. 4 (New York: Wiley Interscience)

Conway A W 1937 *Proc. R. Soc.* **162** A 145–54

—— 1948 *Acta Pontif. Acad. Sci.* **12** 259–77.

Dirac P A M 1945 *Proc. R. Irish Acad.* A **50** 261–70

Edmonds J D 1972 *Int. J. Theor. Phys.* **6** 205–24.

—— 1978 *Found. Phys.* **8** 439–44

Gough W 1984 *Eur. J. Phys.* **5** 163–71

—— 1986 *Eur. J. Phys.* **7** 35–42

—— 1987 *Eur. J. Phys.* **8** 164–70

Kaiser H, George E A and Werner S A 1984 *Phys. Rev.* **29** A 2276–9

Lorrain P and Corson D R 1970 *Electromagnetic Fields and Waves* 2nd edn (San Francisco: Freeman)

Rastall P 1964 *Rev. Mod. Phys.* **36** 820–32

Schiff L I 1955 *Quantum Mechanics* 2nd edn (New York: McGraw-Hill)

Silberstein L 1912 *Phil. Mag.* (6th ser.) **23** 790–809

—— 1913 *Phil. Mag.* (6th ser.) **25** 135–44

Synge J L 1972 *Commun. Dublin Inst. Adv. Studies* **A21** 1–67