The Clifford algebra of physical space and Dirac theory

2016 Eur. J. Phys. 37 055407

(http://iopscience.iop.org/0143-0807/37/5/055407)

IOP Publishing European Journal of Physics

Eur. J. Phys. **37** (2016) 055407 (28pp) doi:10.1088/0143-0807/37/5/055407

The Clifford algebra of physical space and Dirac theory

Jayme Vaz Jr

Departamento de Matemática Aplicada, IMECC Universidade Estadual de Campinas (UNICAMP), 13081-970 Campinas, SP, Brazil

E-mail: vaz@ime.unicamp.br

Received 19 April 2016, revised 20 May 2016
Accepted for publication 9 June 2016
Published 11 July 2016

Abstract

The claim found in many textbooks that the Dirac equation cannot be written solely in terms of Pauli matrices is shown to not be completely true. It is only true as long as the term $\beta\psi$ in the usual Dirac factorization of the Klein–Gordon equation is assumed to be the product of a square matrix β and a column matrix ψ. In this paper we show that there is another possibility besides this matrix product, in fact a possibility involving a matrix operation, and show that it leads to another possible expression for the Dirac equation. We show that, behind this other possible factorization is the formalism of the Clifford algebra of physical space. We exploit this fact, and discuss several different aspects of Dirac theory using this formalism. In particular, we show that there are four different possible sets of definitions for the parity, time reversal, and charge conjugation operations for the Dirac equation.

Keywords: Clifford algebra, Dirac equation, spinors

1. Introduction

In the beginning of quantum mechanics, the explanation of the splitting of a spectral line into several components in the presence of a magnetic field (the Zeeman effect) posed an additional challenge to the brilliant minds that were trying to construct the quantum theory. Among those scientists, Pauli made a remarkable contribution to understanding of the Zeeman effect. In 1924, reasoning in the opposite direction of the models developed so far [1], Pauli proposed that the origin of the multiplicity of a spectral term is the radiant electron itself, instead of the atomic core [2]. Then, in 1925, Uhlenbeck and Goudsmit published the idea of the spin as a new degree of freedom associated with a self-rotating electron [3].

Meanwhile, quantum mechanics emerged in the form of matrix mechanics by Heisenberg, and in the form of wave mechanics by Schrödinger, where the dynamics is described by a $\mathbb{C}$-valued function (the wave function). In 1926, Heisenberg and Jordan applied the new matrix mechanics to the spin model [4], and then Pauli started to study how to incorporate the spin in Schrödinger's wave mechanics. Pauli published his results in 1927 [5]. In order to incorporate the spin in wave mechanics, Pauli used the wave function $\psi(\boldsymbol{x}, t, S_z)$, where S_z is a variable which takes values $\pm\frac{1}{2}$ (in $\hbar$ units). In other words, this means that the wave function incorporating spin has values in a two-dimensional vector space over $\mathbb{C}$, which we denote by $\mathcal{S}$.

In order to construct the Hamiltonian operator, Pauli noticed that, in matrix mechanics, angular momentum matrices satisfy

$$M_a M_b - M_b M_a = \mathrm{i} M_c,$$

where $(a, b, c) = \{(x, y, z), (y, z, x), (z, x, y)\}$ and $\mathrm{i} = \sqrt{-1}$, with

$$|M|^2 = M_x^2 + M_y^2 + M_z^2 = m(m+1),$$

for $m = 0, 1, 2, \ldots$ or $m = 1/2, 3/2, \ldots$, and required that the spin angular momentum operators (s_x, s_y, s_z) satisfy the same relations, that is,

$$s_a s_b - s_b s_a = \mathrm{i} s_c, \tag{1}$$

but with

$$|s|^2 = \frac{1}{2}\left(\frac{1}{2} + 1\right) = \frac{3}{4}.$$

But what are these spin angular momentum operators (s_x, s_y, s_z)?

In order to solve that question, Pauli introduced matrices $(\sigma_1, \sigma_2, \sigma_3)$ given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2}$$

and defined (s_x, s_y, s_z) as

$$s_x = \frac{1}{2}\sigma_1, \quad s_y = \frac{1}{2}\sigma_2, \quad s_z = \frac{1}{2}\sigma_3. \tag{3}$$

Then, Pauli considered $\psi\left(\boldsymbol{x}, t, \frac{1}{2}\right)$ and $\psi\left(\boldsymbol{x}, t, -\frac{1}{2}\right)$ as the components of a two-component vector

$$|\psi(\boldsymbol{x}, t)\rangle = \begin{pmatrix} \psi\left(\boldsymbol{x}, t, \frac{1}{2}\right) \\ \psi\left(\boldsymbol{x}, t, -\frac{1}{2}\right) \end{pmatrix}, \tag{4}$$

writing the wave equation as

$$(\mathbf{H} - \mathrm{i}\hbar\partial_t)|\psi(\boldsymbol{x}, t)\rangle = 0, \tag{5}$$

where the Hamiltonian $\mathbf{H}$ for the motion of an electron in an electromagnetic field described by the vector potential $\mathbf{A} = (A_x, A_y, A_z)$ and scalar potential ϕ is given by

$$\mathbf{H} = \frac{1}{2m}[\sigma_1(p_x - eA_x) + \sigma_2(p_y - eA_y) + \sigma_3(p_z - eA_z)]^2 + e\phi, \tag{6}$$

where $p_a = -\mathrm{i}\hbar\partial_a$ with $a = x, y, z$, and we have taken $c = 1$. Therefore, Pauli made the identification $\mathcal{S} = \mathbb{C}^2$.

Pauli was aware that his theory had to be complemented in order to include relativistic effects, particularly to obtain the correct gyromagnetic factor of the electron. Schrödinger was already studying a relativistic version of wave mechanics using the Klein–Gordon equation, but the results were incompatible with experimental data. It was Dirac, in 1928, who solved this problem [6].

In a standard approach to the Dirac equation—see, for example, [7, 8]—and following Dirac's original derivation, one starts with a first-order equation for $|\psi(\boldsymbol{x}, t)\rangle \in \mathbb{C}^N$ of the form (Einstein summation convection is assumed throughout this paper)

$$\mathrm{i}\hbar(\partial_t + \alpha^k \partial_k)|\psi(\boldsymbol{x}, t)\rangle = m\beta|\psi(\boldsymbol{x}, t)\rangle, \tag{7}$$

where α^k ($k = 1, 2, 3$) and β are supposed to be $N \times N$ matrices over a field $\mathbb{K}$, and in order for this equation to be squared to the Klein–Gordon equation for each of the components of $|\psi(\boldsymbol{x}, t)\rangle$, the quantities α^k ($k = 1, 2, 3$) and β have to satisfy

$$\alpha^j \alpha^k + \alpha^k \alpha^j = 2\delta^{ij} I, \tag{8}$$

$$\alpha^j \beta + \beta \alpha^j = 0, \quad \beta^2 = I, \tag{9}$$

for $j, k = 1, 2, 3$, where I is the identity matrix. As well-known, the smallest number N for which the above relations are satisfied is $N = 4$, and in this case $\mathbb{K} = \mathbb{C}$ and $|\psi(\boldsymbol{x}, t)\rangle$ is a four-component complex-valued column vector.

The use of a $\mathbb{C}^4$-valued wave function is clearly a generalization of the $\mathbb{C}^2$-valued wave function used by Pauli. But is this the only possibility for the wave function space? Since it is reasonable to suppose that Dirac looked for a $\mathbb{C}^N$-valued wave function as a generalization of Pauli's $\mathbb{C}^2$-valued wave function, we might ask first if there is another choice for the spinor space $\mathcal{S}$ besides $\mathbb{C}^2$.

In his 1972 Gibbs lecture, Dyson [9] enumerated some situations where the progress of both mathematics and physics has been retarded by the lack of mutual communication; he referred to these situations as 'missed opportunities'. One of these missed opportunities described by Dyson is the role played by the quaternions in the development of physics and mathematics. The quaternions $\mathbb{H}$ are defined by the set of elements $\{a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} |\, a, b, c, d \in \mathbb{R}\}$ with associative products defined by $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$, $\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$, $\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$, and $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$. However, it is well-known that the quaternion units $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ can be written in terms of Pauli matrices as

$$\boldsymbol{i} = -\mathrm{i}\sigma_1, \quad \boldsymbol{j} = -\mathrm{i}\sigma_2, \quad \boldsymbol{k} = -\mathrm{i}\sigma_3. \tag{10}$$

From this, we see that we can define (s_x, s_y, s_z) by

$$s_x = \mathrm{i}\frac{\boldsymbol{i}}{2}, \quad s_y = \mathrm{i}\frac{\boldsymbol{j}}{2}, \quad s_z = \mathrm{i}\frac{\boldsymbol{k}}{2}. \tag{11}$$

Now there is room to speculate what space $\mathcal{S}$ one could choose if (s_x, s_y, s_z) are defined as in equation (11). It is no surprise that one chooses, in this case, $\mathcal{S} = \mathbb{H}$. Indeed, if we are looking for a generalization of a $\mathbb{C}$-valued function, a $\mathbb{H}$-valued function is a natural candidate, and since a quaternion can be seen as a pair of complex numbers, the number of degrees of freedom required to describe the spin space is matched with this choice. As a matter of fact, the Pauli equation can be written in terms of a quaternionic wave function, but this will not be discussed here since it would divert us from the main issue.

A $\mathbb{H}$-valued wave function $\psi = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ can also be written, because of the representation equation (10), as a 2×2 matrix ψ of the form

$$\psi = \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix} = \begin{pmatrix} |\psi\rangle & -\mathrm{i}\sigma_2\,|\psi\rangle^* \end{pmatrix}, \tag{12}$$

where $\alpha = a - \mathrm{i}d$, $\beta = c - \mathrm{i}b$, and $*$ denotes complex conjugation, and the column matrices $|\psi\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ and $-\mathrm{i}\sigma_2\,|\psi\rangle^* = \begin{pmatrix}-\beta^*\\ \alpha^*\end{pmatrix}$ (which will be interpreted later). Note that we are omitting the dependence of the wave function on the space and time variables.

Now let us look at the factorization of the Klein–Gordon equation under this perspective, that is, considering the possibility that the wave function could be written in terms of a matrix, which we will denote by ψ. Note that an arbitrary 2×2 complex matrix ψ has the same number of degrees of freedom as $|\psi\rangle \in \mathbb{C}^4$. Then, in this case, the matrix multiplication $\beta|\psi\rangle$ in equation (7) is replaced by $\beta\psi$. But the matrix multiplication $\beta\psi$ can be seen as an operation in the space of ψ, that is, we have an operator $\boldsymbol{\beta}$ given by $\boldsymbol{\beta}(\psi) = \beta\psi$. In this case, equation (7) is replaced by

$$\mathrm{i}\hbar(\partial_t + \alpha^k\partial_k)\psi = m\boldsymbol{\beta}(\psi). \tag{13}$$

Let us consider now the possibility that $\boldsymbol{\beta}$ could be an arbitrary $\mathbb{R}$-linear operation on the space of the wave function. In order to square equation (13) to the Klein–Gordon equation, the matrices α^k ($k = 1, 2, 3$) must satisfy the same relation in equation (8), and for $\boldsymbol{\beta}$ we must have

$$\boldsymbol{\beta}(\boldsymbol{\alpha}^i\psi) + \boldsymbol{\alpha}^i\boldsymbol{\beta}(\psi) = 0, \quad i = 1, 2, 3, \tag{14}$$

$$(\mathrm{i}\boldsymbol{\beta})^2(\psi) = \mathrm{i}\boldsymbol{\beta}(\mathrm{i}\boldsymbol{\beta}(\psi)) = -\psi. \tag{15}$$

The matrices α^k can be chosen as the Pauli matrices, but $\boldsymbol{\beta}(\psi)$ cannot be taken as $\beta\psi$ because in this case the relation $\beta\alpha^i + \alpha^i\beta = 0$ ($i = 1, 2, 3$) cannot be satisfied, since the maximum number of (linearly independent and mutually anticommuting) 2×2 complex matrices is three. So, let us keep the choice of α^k as Pauli matrices, and look for a matrix operation such that equations (14) and (15) hold. From equation (8), we have that

$$\boldsymbol{\beta}(\alpha^i\alpha^j) + \boldsymbol{\beta}(\alpha^j\alpha^i) = 2\delta^{ij}\boldsymbol{\beta}(I).$$

But, from equation (14), using $\psi = \alpha^j$,

$$\boldsymbol{\beta}(\alpha^i\alpha^j) = -\alpha^j\boldsymbol{\beta}(\alpha^j).$$

Then, we have

$$\alpha^i\boldsymbol{\beta}(\alpha^j) + \alpha^j\boldsymbol{\beta}(\alpha^i) = -2\delta^{ij}\boldsymbol{\beta}(I),$$

which is satisfied for

$$\boldsymbol{\beta}(\alpha^i) = -\alpha^i\boldsymbol{\beta}(I), \quad i = 1, 2, 3. \tag{16}$$

Using equation (14) again, we have

$$\boldsymbol{\beta}(\alpha^i\psi) = -\alpha^i\boldsymbol{\beta}(\psi) = -\alpha^i\boldsymbol{\beta}(I)\boldsymbol{\beta}^{-1}(I)\boldsymbol{\beta}(\psi) = \boldsymbol{\beta}(\alpha^i)\boldsymbol{\beta}^{-1}(I)\boldsymbol{\beta}(\psi),$$

and from linearity, we conclude that $\boldsymbol{\beta}$ must satisfy

$$\boldsymbol{\beta}(\phi\psi) = \boldsymbol{\beta}(\phi)\boldsymbol{\beta}^{-1}(I)\boldsymbol{\beta}(\psi). \tag{17}$$

Now, using equations (15), (16) and (17), we have

$$-\alpha^j = (\mathrm{i}\boldsymbol{\beta})^2(\alpha^j) = -\mathrm{i}\boldsymbol{\beta}(\mathrm{i}\alpha^j\boldsymbol{\beta}(I)) = -\mathrm{i}\boldsymbol{\beta}(\mathrm{i}\alpha^j)\boldsymbol{\beta}^{-1}(I)\boldsymbol{\beta}^2(I).$$

Let us also write

$$\boldsymbol{\beta}(\mathrm{i}\psi) = \epsilon\mathrm{i}\boldsymbol{\beta}(\psi),$$

that is, $\boldsymbol{\beta}$ is $\mathbb{C}$-linear or $\mathbb{C}$-antilinear according to $\epsilon = 1$ or $\epsilon = -1$, respectively. Then, we have

$$-\alpha^j = -\epsilon\alpha^j\boldsymbol{\beta}^2(I),$$

and consequently

$$\boldsymbol{\beta}^2(I) = \epsilon.$$

But, since the matrices α^k are chosen as the Pauli matrices, we have relations like $\alpha^1\alpha^2 = \mathrm{i}\alpha^3$. Let us see the effect of $\boldsymbol{\beta}$ on this relation; on the LHS, we have

$$\begin{aligned}\boldsymbol{\beta}(\alpha^1\alpha^2) &= \boldsymbol{\beta}(\alpha^1)\boldsymbol{\beta}^{-1}(I)\boldsymbol{\beta}(\alpha^2) = (-\alpha^1)\boldsymbol{\beta}(I)\boldsymbol{\beta}^{-1}(I)(-\alpha^2)\boldsymbol{\beta}(I)\\ &= \alpha^1\alpha^2\boldsymbol{\beta}(I) = \mathrm{i}\alpha^3\boldsymbol{\beta}(I),\end{aligned}$$

and on the RHS,

$$\boldsymbol{\beta}(\mathrm{i}\alpha^3) = \epsilon\mathrm{i}\boldsymbol{\beta}(I)(\alpha^3) = -\epsilon\mathrm{i}\alpha^3\boldsymbol{\beta}(I),$$

and comparing both expressions, we conclude that we must have $\epsilon = -1$. Thus, $\boldsymbol{\beta}$ is such that

$$\boldsymbol{\beta}(\mathrm{i}\psi) = -\mathrm{i}\boldsymbol{\beta}(\psi). \tag{18}$$

Using equations (16) and (18) we can conclude, after writing an arbitrary 2×2 complex matrix as a linear combination of the Pauli matrices and the identity matrix, that

$$\boldsymbol{\beta}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d^* & -c^* \\ -b^* & a^* \end{pmatrix}\boldsymbol{\beta}(I), \tag{19}$$

and from equations (15) and (18) that

$$\boldsymbol{\beta}^2 = -I. \tag{20}$$

$\boldsymbol{\beta}(I)$ remains to be found. Let us suppose, for simplicity, that it is a diagonal matrix, $\boldsymbol{\beta}(I) = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$. From equations (19) and (20), we conclude that we must have $xy^* = -1$. Therefore, a simple solution is

$$\boldsymbol{\beta}(I) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3, \tag{21}$$

and then

$$\boldsymbol{\beta}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d^* & c^* \\ -b^* & -a^* \end{pmatrix}. \tag{22}$$

Note that we can also write

$$\boldsymbol{\beta}(\psi) = \mathrm{adj}(\psi^\dagger)\sigma_3, \tag{23}$$

where $\mathrm{adj}(\psi^\dagger)$ denotes the classical adjoint (adjugate) matrix of $\psi^\dagger$ and $\dagger$ denotes a Hermitian conjugation.

We see, therefore, that it is possible to write the Dirac equation using only Pauli matrices, as long as the fourth matrix needed in the factorization process is replaced by a matrix operation, which is the one defined in equation (23). This new form of Dirac equation is given in equation (13). The matrix ψ is a 2×2 complex matrix, which can be written, in analogy to equation (12), as

$$\psi = \begin{pmatrix} |\xi\rangle & -\mathrm{i}\sigma_2\,|\eta\rangle^* \end{pmatrix}. \tag{24}$$

It is interesting to see the role played by $|\xi\rangle$ and $|\eta\rangle$. In fact, using equation (23), we have

$$\boldsymbol{\beta}(\psi) = \begin{pmatrix} |\eta\rangle & \mathrm{i}\sigma_2\,|\xi\rangle^* \end{pmatrix}, \tag{25}$$

and from equation (13) we obtain

$$i\hbar(\partial_t + \alpha^k\partial_k)|\xi\rangle = m|\eta\rangle \tag{26}$$

and

$$i\hbar(\partial_t + \alpha^k\partial_k)(-i\sigma_2)|\eta\rangle^* = m(i\sigma_2)|\xi\rangle^*. \tag{27}$$

But $(\alpha^1)^* = \alpha^1$, $(\alpha^2)^* = -\alpha^2$ and $(\alpha^3)^* = \alpha^3$, and it follows that

$$i\hbar(\partial_t - \alpha^k\partial_k)|\eta\rangle = m|\xi\rangle. \tag{28}$$

This suggests that $|\xi\rangle$ and $|\eta\rangle$ can be interpreted as left and right spinors—and we will see later, after analyzing how they transform under a Lorentz transformation, that this is indeed the case. However, unlike in the usual factorization process, where left and right spinors are combined in the form of a column matrix, they have appeared combined in the form of a square matrix!

What have we missed for not using quaternions, or for not thinking in terms of quaternions, in some moment in the development of quantum mechanics? The mathematical structure behind the above discussion concerning the factorization of the Klein–Gordon equation using 2×2 complex matrices is called *the Clifford algebra of physical space*, or simply the *algebra of physical space* (APS) [10]. This algebra of 2×2 complex matrices $\mathcal{M}(2, \mathbb{C})$ is isomorphic to the algebra of complexified quaternions $\mathbb{C} \otimes \mathbb{H}$. The interesting fact here is that $\mathbb{C}$, $\mathbb{H}$ and $\mathcal{M}(2, \mathbb{C}) \simeq \mathbb{C} \otimes \mathbb{H}$ are all examples of a Clifford algebra, and one is a particular Clifford subalgebra (the even subalgebra) of another, that is, if we denote the even subalgebra by a plus sign as superscript, we have $\mathbb{C} \simeq \mathbb{H}^+$ and $\mathbb{H} \simeq (\mathbb{C} \otimes \mathbb{H})^+$ [11–13].

But, is there any advantage in using 2×2 complex matrices instead of $\mathbb{C}^4$-valued column matrices in Dirac theory? We believe that there is, and we classify the advantages as computational, didactical and epistemological. The computational advantages are seen, for instance, when we notice that, in many cases, a square matrix possesses an inverse matrix, whereas a column matrix does not. The existence of an inverse element makes it easier to manipulate some mathematical expressions, and the proof of Fierz identities is a very good example of this computational advantage, as we will see (in section 4). The didactical advantages are manifested by the fact that the same mathematical structure that can be used to study mechanics, in particular rigid body kinematics (in terms of the Cayley–Klein parameters) [14], and electromagnetism (see appendix) can be used to study quantum mechanics. In other words, there is no need for an additional mathematical structure in relativistic quantum mechanics besides the one already used in classical mechanics and electromagnetism. In order to grasp the epistemological advantages, we must take into account the fact that the 2×2 complex matrix algebra is in fact a *representation* of an algebra constructed from entities with a clear geometrical meaning. This is the APS. The elements of this algebra are the representatives of geometrical objects that are oriented line segments, oriented plane fragments, and oriented volumes. For this reason, the original denomination given by Clifford for this mathematical structure was *geometric algebra*.

The main objective of this paper is to present Dirac theory formulated exclusively in terms of the APS. The use of Clifford algebras in Dirac theory has been extensively discussed by Hestenes (see, for example, [15, 16]), but Hestenes' approach is based on the so-called *spacetime algebra*, that is, the Clifford algebra of Minkowski spacetime. The spacetime algebra is not the Dirac algebra of gamma matrices usually introduced in quantum mechanics textbooks; indeed, the Dirac algebra is the *complexification* of the spacetime algebra. But, while the spacetime algebra has half of the dimension of the Dirac algebra, it is still too large for the sake of formulating Dirac theory, since the APS has half of the dimension of the spacetime algebra. In this sense, we can say that the approach to Dirac theory using the APS is a *minimalistic* one, that is, the APS is the smaller algebraic structure suitable for the formulation of Dirac theory.

Applications of the APS in physics have been widely discussed [17–25], and we aim here to complement the presentations concerning relativistic quantum mechanics. This paper is organized as follows. In section 2 we introduce the APS and discuss its main properties. We discuss its matrix representation and its relation to the $\boldsymbol{\beta}$ operation previously discussed. In section 3 we introduce the concept of paravectors. The need for this concept comes from the fact that the APS is an algebra for a three-dimensional space, and we need to describe vectors in a four-dimensional space (the spacetime). We also discuss how to deal with Lorentz transformations in this formalism, as well as its matrix representation. In section 4 we introduce the Dirac equation with an electromagnetic coupling. Using the APS, we discuss Weyl spinors, the Lorentz covariance of the APS form of the Dirac equation, the bilinear covariants and Fierz identities. In section 5 we prove some conservation laws associated with the Dirac equation. Finally, in section 6 we discuss the effect of parity, time reversal and charge conjugation operations in the Dirac equation using the APS. We show that we have four different possible sets of definitions for these operations.

2. The Clifford APS

The real Clifford algebra $\mathcal{C}\ell_3$ of the Euclidean 3-space $(\mathbb{R}^3, g)$ is the associative algebra generated by $\{\mathbf{v}|\mathbf{v} \in \mathbb{R}^3\}$ and $\{a1|a \in \mathbb{R}\}$ subject to the condition [11–13]

$$\mathbf{vu} + \mathbf{uv} = 2g(\mathbf{v}, \mathbf{u}), \tag{29}$$

where $g(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$ for $\{\mathbf{e}_i\}$ $(i = 1, 2, 3)$ being the canonical basis of $\mathbb{R}^3$. We call this algebra the APS. An arbitrary element of $\mathcal{C}\ell_3$ is a sum of elements of the form

$$\mathbf{e}_1^{\mu_1}\mathbf{e}_2^{\mu_2}\mathbf{e}_3^{\mu_3}, \tag{30}$$

with $\mu_i = 0$ or 1 $(i = 1, 2, 3)$, and the case $\mu_1 = \mu_2 = \mu_3 = 0$ understood as the unit. In this case there is a Z_n-gradation and we call an element in equation (30) a k-vector according to $\mu_1 + \mu_2 + \mu_3 = k$. A 0-vector is, therefore, a scalar; a 1-vector is a vector associated with a class of equipollent line segments; a 2-vector is a vector associated with the equivalence class of oriented plane fragments with the same area and direction; and a 3-vector is a vector associated with the equivalence class of volume elements with the same orientation and volume. The 3-vector

$$\boldsymbol{i} = \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3 = \mathbf{e}_{123} \tag{31}$$

deserves special attention since it belongs to the center of $\mathcal{C}\ell_3$ and $\boldsymbol{i}^2 = -1$; it is called the *pseudo-scalar* of $\mathcal{C}\ell_3$. We denote the vector space of k-vectors by $\bigwedge_k(\mathbb{R}^3) = \bigwedge_k$ and adopt the convention that $\bigwedge_0 = \mathbb{R}$ and $\bigwedge_1 = \mathbb{R}^3$. An arbitrary element of $\bigwedge = \oplus_{j=0}^{3}\bigwedge_j$ is called a *multivector*. The projectors $\bigwedge \to \bigwedge_k$ are denoted by $\langle\ \rangle_k$, that is, the projector $\langle\ \rangle_k$ extracts the k-vector part of a multivector. We also use the notation $\mathbf{e}_{ij} = \mathbf{e}_i\mathbf{e}_j$ for $i \neq j$.

The Clifford product of a vector and an arbitrary k-vector A_k can be decomposed as

$$\mathbf{v}A_k = \mathbf{v}\cdot A_k + \mathbf{v}\wedge A_k, \tag{32}$$

where the interior and exterior products are defined, respectively, as

$$\mathbf{v}\cdot A_k = \langle \mathbf{v}A_k\rangle_{k-1} = \frac{1}{2}(\mathbf{v}A_k - (-1)^k A_k\mathbf{v}),$$
$$\mathbf{v}\wedge A_k = \langle \mathbf{v}A_k\rangle_{k+1} = \frac{1}{2}(\mathbf{v}A_k + (-1)^k A_k\mathbf{v}). \tag{33}$$

We can also define three operations in $\mathcal{C}\ell_3$ called grade involution (or parity, denoted by a hat), reversion (denoted by a tilde) and (Clifford) conjugation (denoted by a bar), respectively as

$$\hat{A}_k = (-1)^k A_k, \qquad \tilde{A}_k = (-1)^{k(k-1)/2} A_k, \qquad \bar{A}_k = \hat{\tilde{A}}_k = \tilde{\hat{A}}_k. \tag{34}$$

The name reversion operation comes from the property where

$$\widetilde{AB} = \tilde{B}\tilde{A}. \tag{35}$$

Conjugation is the composition of grade involution and reversion. From its definition, grade involution satisfies

$$\widehat{AB} = \hat{A}\hat{B}. \tag{36}$$

Therefore, conjugation satisfies

$$\overline{AB} = \bar{B}\bar{A}. \tag{37}$$

The following table summarizes the effect of these operations on k-vectors:

k	$\hat{}$	$\tilde{}$	$\bar{}$
0	+	+	+
1	−	+	−
2	+	−	−
3	−	−	+

Elements of $\mathcal{C}\ell_3$ are called even if $A = \hat{A}$ and odd if $A = -\hat{A}$. The set of even elements is closed under the Clifford product and is called the even subalgebra $\mathcal{C}\ell_3^+$. A very important feature of Clifford algebras is that even subalgebras are also Clifford algebras, and in this case we have

$$\mathcal{C}\ell_3^+ \simeq \mathbb{H}. \tag{38}$$

Another distinguished feature is that the vector space of 2-vectors is isomorphic to the vector space of 1-vectors; this isomorphism is provided by the Hodge duality $\star$, which is given by $\star A = \boldsymbol{i}\tilde{A}$, such that $\boldsymbol{i}A_k \in \bigwedge_{3-k}$.

Finally, a useful property is that [11–13]

$$\langle AB\rangle_0 = \langle BA\rangle_0, \tag{39}$$

whose interpretation will be clear in what follows.

The nabla operator ∇ is defined as

$$\nabla = \mathbf{e}^i\partial_i, \tag{40}$$

where $\partial_i = \frac{\partial}{\partial x^i}$ and $\mathbf{e}^i = \mathbf{e}_i$ for $i = 1, 2, 3$ as the Cartesian coordinate basis. From equation (32), we can decompose the action of ∇ on a k-vector A_k as

$$\nabla A_k = \nabla\cdot A_k + \nabla\wedge A_k. \tag{41}$$

When we have a vector $\mathbf{v}$, this results in

$$\nabla\mathbf{v} = \operatorname{div}\mathbf{v} + \boldsymbol{i}\ \operatorname{rot}\mathbf{v}, \tag{42}$$

where

$$\operatorname{div}\mathbf{v} = \nabla\cdot\mathbf{v}, \tag{43}$$

and

$$\operatorname{rot}\mathbf{v} = \nabla\times\mathbf{v} = -\boldsymbol{i}\nabla\wedge\mathbf{v}, \tag{44}$$

where in the last equality we identify the vector product with the Hodge dual of the exterior product [11, 13], that is, given vectors $\mathbf{u}$ and $\mathbf{v}$,

$$\mathbf{u}\times\mathbf{v} = -\boldsymbol{i}(\mathbf{u}\wedge\mathbf{v}). \tag{45}$$

2.1. Matrix representation

Let us consider the matrix representation of $\mathcal{C}\ell_3$. From the theory of the representations of Clifford algebras, we know that $\mathcal{C}\ell_3 \simeq \mathcal{M}(2, \mathbb{C})$. If we represent the vectors $\mathbf{e}_i$ by the Pauli matrices σ_i $(i = 1, 2, 3)$, we obtain that an arbitrary element $A \in \mathcal{C}\ell_3$, that is,

$$A = a_0 + a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 + a_{12}\mathbf{e}_{12} + a_{13}\mathbf{e}_{13} + a_{23}\mathbf{e}_{23} + a_{123}\boldsymbol{i} \tag{46}$$

is represented by the matrix $\mathsf{A} = \operatorname{rep}(A)$ given by

$$\mathsf{A} = \begin{pmatrix} z_1 & z_3 \\ z_2 & z_4 \end{pmatrix} \tag{47}$$

where

$$z_1 = (a_0 + a_3) + \mathrm{i}(a_{12} + a_{123}), \qquad z_2 = (a_1 + a_{13}) + \mathrm{i}(a_2 + a_{23}),$$
$$z_3 = (a_1 - a_{13}) - \mathrm{i}(a_2 - a_{23}), \qquad z_4 = (a_0 - a_3) - \mathrm{i}(a_{12} - a_{123}).$$

When this is restricted to $A_+ \in \mathcal{C}\ell_3^+$ we have

$$\mathsf{A}_+ = \begin{pmatrix} w_1 & -w_2^* \\ w_2 & w_1^* \end{pmatrix} \tag{48}$$

where

$$w_1 = a_0 + \mathrm{i}a_{12}, \quad w_2 = a_{13} + \mathrm{i}a_{23}.$$

Note that this is exactly the same type of matrix as in equation (12), and it is a consequence of the fact that $\mathcal{C}\ell_3^+ \simeq \mathbb{H}$. Note also that

$$\langle A\rangle_0 = \frac{1}{2}\operatorname{Re}\operatorname{Tr}(\mathsf{A}). \tag{49}$$

This result reflects in a property like the one in equation (39).

The operations of grade involution $\hat{A}$ and reversion $\tilde{A}$ are expressed in this matrix representation as

$$\operatorname{rep}(\hat{A}) = \operatorname{adj}(\mathsf{A}^\dagger) = \begin{pmatrix} z_4^* & -z_2^* \\ -z_3^* & z_1^* \end{pmatrix}, \qquad \operatorname{rep}(\tilde{A}) = \mathsf{A}^\dagger = \begin{pmatrix} z_1^* & z_2^* \\ z_3^* & z_4^* \end{pmatrix}. \tag{50}$$

If we compare $\mathrm{rep}(\hat{A})$ with equation (23) we see that

$$\boldsymbol{\beta}(\psi) = \mathrm{rep}(\hat{\psi}\mathbf{e}_3). \tag{51}$$

3. Paravectors and Lorentz transformations

Four-dimensional (spacetime) vectors can be described in the APS in terms of *paravectors*, that is, elements of $\bigwedge_0 \oplus \bigwedge_1$. Important examples are momentum $p = mv$, electromagnetic potential A, proper velocity v and spin s, that is,

$$p = E + \mathbf{p}, \qquad A = \phi + \mathbf{A}, \qquad v = v_0 + \mathbf{v}, \qquad s = s_0 + \mathbf{s}, \tag{52}$$

where we use $c = 1$ and E is the energy, $\mathbf{p}$ is the momentum, ϕ is the scalar potential, $\mathbf{A}$ is the vector potential, and, for example, for a particle in the rest frame, $v_0 = 1$, $\mathbf{v} = 0$, $s_0 = 0$ and $\mathbf{s} = \frac{1}{2}\mathbf{e}_3$. Paravectors can also be defined as elements a of the APS that satisfy $a = \tilde{a}$.

Four-dimensional scalar products can be defined as

$$u \bullet v = \frac{1}{2}(u\bar{v} + v\bar{u}). \tag{53}$$

If we write $u = u_0 + \mathbf{u}$ and $v = v_0 + \mathbf{v}$, it is easy to see that

$$u \bullet v = u_0 v_0 - \mathbf{u} \cdot \mathbf{v}. \tag{54}$$

For the momentum p we have

$$|p|^2 = p\bar{p} = E^2 - \mathbf{p}^2 = m^2. \tag{55}$$

Four-dimensional bivectors can be described in the APS in terms of *biparavectors*, that is, elements of $\bigwedge_1 \oplus \bigwedge_2$. Important examples are the electromagnetic field F and the intrinsic electromagnetic moment S,

$$F = \mathbf{E} + \boldsymbol{i}\mathbf{B}, \qquad S = \mathbf{d} + \boldsymbol{i}\mathbf{m}, \tag{56}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively, and $\mathbf{d}$ and $\mathbf{m}$ are the electric and magnetic moments, respectively. Biparavectors can also be defined as elements B of the APS that satisfy $B = -\bar{B}$.

The four-dimensional wedge product can be written in the APS as

$$u \mathbin{\bar{\wedge}} v = \frac{1}{2}(u\bar{v} - v\bar{u}). \tag{57}$$

Note that

$$\begin{aligned} u \mathbin{\bar{\wedge}} v &= v_0\mathbf{u} - u_0\mathbf{v} - \mathbf{u} \wedge \mathbf{v} \\ &= v_0\mathbf{u} - u_0\mathbf{v} + \boldsymbol{i}\mathbf{u} \times \mathbf{v}, \end{aligned} \tag{58}$$

where $\mathbf{u} \times \mathbf{v} = -\boldsymbol{i}\mathbf{u} \wedge \mathbf{v}$ is the vector product [11, 13].

Now suppose it is given a scalar quantity. Then the question arises: is this scalar quantity a pure scalar or is it part of a paravector? In the same way, given a vector, how can we identify it as part of a paravector or as part of a biparavector? The answer to these questions comes from the behavior under Lorentz transformations. A pure scalar quantity does not change under a Lorentz transformation, but a scalar quantity that is part of a paravector changes; on the other hand, a vector can be distinguished as part of a paravector or as part of a biparavector according to its different transformation rules in each case.

The Lorentz transformation of a paravector p can be written as

$$p \mapsto p' = Rp\tilde{R}, \tag{59}$$

where R is such that

$$R\bar{R} = \bar{R}R = 1. \tag{60}$$

We can write R in the form [11]

$$R = \pm e^{\Omega/2}, \tag{61}$$

with

$$\Omega = \boldsymbol{\nu} + i\boldsymbol{\theta}. \tag{62}$$

When $\boldsymbol{\nu} = 0$ this element describes through equation (59) a rotation in the plane with a normal vector $\boldsymbol{\theta}/\theta$ by the angle $\theta = |\boldsymbol{\theta}|$, whereas when $\boldsymbol{\theta} = 0$ it describes a boost in the direction of the relative velocity $\mathbf{v}$ of the reference frame, such that $\boldsymbol{\nu}/\nu = \mathbf{v}/v$, and with rapidity $\nu = |\boldsymbol{\nu}|$. A Lorentz transformation described by R can be written in terms of the composition of a rotation U and a Lorentz boost L as

$$R = LU, \tag{63}$$

where L and U satisfy $L = \bar{L}^{-1} = \tilde{L}$ and $U = \bar{U}^{-1} = \hat{U}$, respectively.

The Lorentz transformation of a biparavector B is

$$B \mapsto B' = RB\bar{R}, \tag{64}$$

with R as in equation (60). It is important to note the difference in the transformation rules in equation (59) and in equation (64); therefore, given a vector quantity, it is part of a paravector or of a biparavector if it transforms under a Lorentz transformation as in equation (59) or as as in equation (64), respectively.

3.1. Matrix representation

Let us consider, for example, the matrix representation of the paravector $p = E + \mathbf{p}$. From equation (47), we have

$$\mathrm{rep}(p) = \begin{pmatrix} E + p_3 & p_1 - ip_2 \\ p_1 + ip_2 & E - p_3 \end{pmatrix} = (\mathrm{rep}(p))^\dagger, \tag{65}$$

that is, paravectors are represented by Hermitian matrices.

Lorentz transformations satisfy equation (60), that is, $R^{-1} = \bar{R}$. In terms of matrices, this equation reads

$$\mathrm{rep}(R^{-1}) = \mathsf{R}^{-1} = \mathrm{adj}(\mathsf{R}) = \mathrm{rep}(\bar{R}), \tag{66}$$

where we use the fact that the matrix representation of $\bar{R}$ is the composition of the operations in equation (50), since $\bar{R} = \hat{\tilde{R}} = \tilde{\hat{R}}$. But, since $\mathrm{adj}(\mathsf{A}) = \det(\mathsf{A})\mathsf{A}^{-1}$ for an arbitrary invertible matrix A, we conclude that

$$\det(\mathsf{R}) = 1, \tag{67}$$

that is, $\mathsf{R} = \mathrm{rep}(R) \in \mathrm{SL}(2, \mathbb{C})$. The matrix representation of R given by equation (61) is

$$\mathsf{R} = e^{\Omega/2}, \qquad \Omega = \nu + i\theta, \tag{68}$$

where $\nu = \mathrm{rep}(\boldsymbol{\nu})$ and $\theta = \mathrm{rep}(\boldsymbol{\theta})$.

4. Dirac equation

The Dirac equation in the APS is [18, 19]

$$\hbar(\partial_t + \nabla)\psi \boldsymbol{i}\mathbf{e}_3 = m\hat{\psi} + e(\phi - \mathbf{A})\psi, \tag{69}$$

where ψ is an element of $\mathcal{C}\ell_3$ (in fact, it is a field, $\psi = \psi(x) = \psi(t, \boldsymbol{x})$), that is,

$$\begin{aligned} \psi &= c_0 + c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + c_3\mathbf{e}_3 + c_{12}\mathbf{e}_{12} + c_{13}\mathbf{e}_{13} + c_{23}\mathbf{e}_{23} + c_{123}\mathbf{e}_{123} \\ &= (c_0 + \boldsymbol{i}c_{123}) + (c_{12} - \boldsymbol{i}c_3)\mathbf{e}_{12} + (c_{13} + \boldsymbol{i}c_2)\mathbf{e}_{13} + (c_{23} - \boldsymbol{i}c_1)\mathbf{e}_{23}, \end{aligned} \tag{70}$$

and ∇ is given by

$$\nabla = \mathbf{e}_1\frac{\partial}{\partial x^1} + \mathbf{e}_2\frac{\partial}{\partial x^2} + \mathbf{e}_3\frac{\partial}{\partial x^3}. \tag{71}$$

The Dirac operator ∂ is

$$\partial = \partial_t - \nabla, \tag{72}$$

and then the operator in equation (69) can be identified with $\bar{\partial}$. Using the Dirac operator, equation (69) can be written in the simple and elegant form

$$\hbar\bar{\partial}\psi \boldsymbol{i}\mathbf{e}_3 = m\hat{\psi} + e\bar{A}\psi, \tag{73}$$

where A is the electromagnetic potential paravector. The four-dimensional Laplacian (d'Alembertian) is

$$\Box = \partial_t^2 - \nabla^2 = \partial\bar{\partial} = \bar{\partial}\partial. \tag{74}$$

The Dirac equation can also be written in the form

$$\boldsymbol{i}\hbar\partial_t\psi\mathbf{e}_3 = \mathsf{H}[\psi] \tag{75}$$

with the Hamiltonian H given by

$$\mathsf{H}[\psi] = -\hbar\nabla\psi \boldsymbol{i}\mathbf{e}_3 + m\hat{\psi} + e(\phi - \mathbf{A})\psi. \tag{76}$$

The momentum operator $\mathbf{p} = \mathbf{e}^i\mathsf{p}_i$ is

$$\mathbf{p}[\psi] = -\hbar\nabla\psi \boldsymbol{i}\mathbf{e}_3. \tag{77}$$

4.1. Weyl spinors

Let $\psi_\pm = \frac{1}{2}(\psi \pm \hat{\psi})$ and

$$f_\pm = \frac{1}{2}(1 \pm \mathbf{e}_3). \tag{78}$$

Using $\mathbf{e}_3 f_\pm = \pm f_\pm$, we can write

$$\begin{aligned} \psi &= (\psi_+ + \psi_-)(f_+ + f_-) \\ &= \psi_+ f_+ + \psi_-\mathbf{e}_3 f_+ + \psi_+ f_- + \psi_-(-\mathbf{e}_3)f_- \\ &= (\psi_+ + \psi_-\mathbf{e}_3)f_+ + (\psi_+ - \psi_-\mathbf{e}_3)f_- \\ &= \xi f_+ + \eta f_-, \end{aligned} \tag{79}$$

where $\xi, \eta \in \mathcal{C}\ell_3^+$. Using this decomposition of ψ in the above form of the Dirac equation we get

$$\boldsymbol{i}\hbar(\partial_t + \nabla)\xi f_+ - \boldsymbol{i}\hbar(\partial_t + \nabla)\eta f_- = m\xi f_- + m\eta f_+ + e(\phi - \mathbf{A})\xi f_+ + e(\phi - \mathbf{A})\eta f_-, \tag{80}$$

which can be decomposed into two equations, that is,

$$\boldsymbol{i}\hbar(\partial_t + \nabla)\xi f_+ = m\eta f_+ + e(\phi - \mathbf{A})\xi f_+, \tag{81}$$

$$-\boldsymbol{i}\hbar(\partial_t + \nabla)\eta f_- = m\xi f_- + e(\phi - \mathbf{A})\eta f_-. \tag{82}$$

The last equation can be transformed into an equation in the algebraic ideal $\mathcal{C}\ell_3 f_+$ with the use of the grade involution, and then

$$\boldsymbol{i}\hbar[(\partial_t + \boldsymbol{i}e\phi) + (\nabla - \boldsymbol{i}e\mathbf{A})]\xi f_+ = m\eta f_+, \tag{83}$$

$$\boldsymbol{i}\hbar[(\partial_t + \boldsymbol{i}e\phi) - (\nabla - \boldsymbol{i}e\mathbf{A})]\eta f_+ = m\xi f_+. \tag{84}$$

In this way we are tempted to identify the left and right components of ψ with ξf_+ and ηf_+, respectively; however, from equation (79) we see that ψ can be written as a sum of elements belonging to two different ideals of $\mathcal{C}\ell_3$, that is, ones of the form $\mathcal{C}\ell_3 f_\pm$, and if we associate the left component of ψ with the ideal $\mathcal{C}\ell_3 f_+$, it is natural to associate the right component of ψ with the ideal $\mathcal{C}\ell_3 f_-$. This situation can be handled with the identification of the left and right components of ψ as

$$\xi f_+ = \psi_{\mathrm{L}}, \qquad \eta f_+ = \psi_{\mathrm{R}}\mathbf{e}_{13}. \tag{85}$$

Note that $\psi_{\mathrm{L}} \in \mathcal{C}\ell_3 f_+$ and $\psi_{\mathrm{R}} \in \mathcal{C}\ell_3 f_-$, such that

$$\boldsymbol{i}\hbar[(\partial_t + \boldsymbol{i}e\phi) + (\nabla - \boldsymbol{i}e\mathbf{A})]\psi_{\mathrm{L}} = m\psi_{\mathrm{R}}\mathbf{e}_{13}, \tag{86}$$

$$\boldsymbol{i}\hbar[(\partial_t + \boldsymbol{i}e\phi) - (\nabla - \boldsymbol{i}e\mathbf{A})]\psi_{\mathrm{R}}\mathbf{e}_{13} = m\psi_{\mathrm{L}}. \tag{87}$$

The spinor field ψ can be written as

$$\psi = \psi_{\mathrm{L}} + \widehat{\psi_{\mathrm{R}}}\mathbf{e}_{13}, \tag{88}$$

with

$$\psi_{\mathrm{L}} = \psi f_+, \qquad \psi_{\mathrm{R}} = \widehat{\psi}\mathbf{e}_{31} f_-. \tag{89}$$

The effect of a Lorentz transformation on ψ (see next subsection) is given by

$$\psi \mapsto \psi' = R\psi = R\psi_{\mathrm{L}} + \widehat{\hat{R}\psi_{\mathrm{R}}}\mathbf{e}_{13}, \tag{90}$$

that is, the left component transforms as

$$\psi_{\mathrm{L}} \mapsto \psi'_{\mathrm{L}} = R\psi_{\mathrm{L}}, \tag{91}$$

and the right component transforms as

$$\psi_{\mathrm{R}} \mapsto \psi'_{\mathrm{R}} = \hat{R}\psi_{\mathrm{R}}. \tag{92}$$

4.1.1. Matrix representation of Weyl spinors. Using the definitions given by equation (89), we can identify the components of ψ with the ones of ψ_{L} and ψ_{R}. Let us denote the matrix representation of these objects by ψ, ψ_{L} and ψ_{R}, respectively. Then

$$\psi = \begin{pmatrix} \boldsymbol{\psi}_{\mathrm{L}} & -\mathrm{i}\sigma_2\boldsymbol{\psi}_{\mathrm{R}}^* \end{pmatrix}, \qquad \psi_{\mathrm{L}} = \begin{pmatrix} \boldsymbol{\psi}_{\mathrm{L}} & 0 \end{pmatrix}, \qquad \psi_{\mathrm{R}} = \begin{pmatrix} 0 & \boldsymbol{\psi}_{\mathrm{R}} \end{pmatrix} \tag{93}$$

where we denote

$$\boldsymbol{\psi}_{\mathrm{L}} = \begin{pmatrix} \psi_{\mathrm{L}1} \\ \psi_{\mathrm{L}2} \end{pmatrix}, \quad \boldsymbol{\psi}_{\mathrm{R}} = \begin{pmatrix} \psi_{\mathrm{R}1} \\ \psi_{\mathrm{R}2} \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{94}$$

and by $*$ the complex conjugation. Note that

$$\hat{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{\mathrm{R}} & -\mathrm{i}\sigma_2 \boldsymbol{\psi}_{\mathrm{L}}^* \end{pmatrix}, \qquad \hat{\psi}_{\mathrm{L}} = \begin{pmatrix} 0 & -\mathrm{i}\sigma_2 \boldsymbol{\psi}_{\mathrm{L}}^* \end{pmatrix}, \qquad \hat{\psi}_{\mathrm{R}} = \begin{pmatrix} \mathrm{i}\sigma_2 \boldsymbol{\psi}_{\mathrm{R}}^* & 0 \end{pmatrix}. \tag{95}$$

In this way, the transformation $\psi \mapsto \psi' = \mathsf{R}\psi$ is

$$\begin{pmatrix} \psi'_{\mathrm{L}1} & -{\psi'_{\mathrm{R}2}}^* \\ \psi'_{\mathrm{L}2} & {\psi'_{\mathrm{R}1}}^* \end{pmatrix} = \begin{pmatrix} r_1 & r_3 \\ r_2 & r_4 \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{L}1} & -\psi_{\mathrm{R}2}^* \\ \psi_{\mathrm{L}2} & \psi_{\mathrm{R}1}^* \end{pmatrix},$$

which gives, for the column vectors,

$$\begin{pmatrix} \psi'_{\mathrm{L}1} \\ \psi'_{\mathrm{L}2} \end{pmatrix} = \begin{pmatrix} r_1 & r_3 \\ r_2 & r_4 \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{L}1} \\ \psi_{\mathrm{L}2} \end{pmatrix}, \qquad \begin{pmatrix} \psi'_{\mathrm{R}1} \\ \psi'_{\mathrm{R}2} \end{pmatrix} = \begin{pmatrix} r_4^* & -r_2^* \\ -r_3^* & r_1^* \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{R}1} \\ \psi_{\mathrm{R}2} \end{pmatrix},$$

that is,

$$\boldsymbol{\psi}_{\mathrm{L}} \mapsto \psi'_{\mathrm{L}} = \mathsf{R}\boldsymbol{\psi}_{\mathrm{L}}, \qquad \boldsymbol{\psi}_{\mathrm{R}} \mapsto \psi'_{\mathrm{R}} = \mathrm{adj}(\mathsf{R}^\dagger)\boldsymbol{\psi}_{\mathrm{R}}. \tag{96}$$

If we use R as in equation (68), it follows that

$$\psi'_{\mathrm{L}} = \mathrm{e}^{(\nu+\mathrm{i}\theta)/2}\boldsymbol{\psi}_{\mathrm{L}}, \qquad \psi'_{\mathrm{R}} = \mathrm{e}^{(-\nu+\mathrm{i}\theta)/2}\boldsymbol{\psi}_{\mathrm{R}}, \tag{97}$$

showing that $\boldsymbol{\psi}_{\mathrm{L}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ transform under different and non-equivalent $\mathrm{SL}(2, \mathbb{C})$ matrix representations of the Lorentz transformation R.

4.2. Lorentz covariance

Let us multiply equation (69) on the left by a Lorentz transformation $\hat{R}$, that is,

$$\hbar\hat{R}(\partial_t + \nabla)\psi\mathbf{ie}_3 = m\hat{R}\hat{\psi} + e\hat{R}(\phi - \mathbf{A})\psi, \tag{98}$$

But, since $R\bar{R} = \bar{R}R = 1$, we can write

$$\hbar\hat{R}(\partial_t + \nabla)\bar{R}R\psi\mathbf{ie}_3 = m\hat{R}\hat{\psi} + e\hat{R}(\phi - \mathbf{A})\bar{R}R\psi. \tag{99}$$

The Dirac operator ∂ and the eletromagnetic potential A transform as in equation (59), that is,

$$\partial' = R\partial\bar{R}, \qquad A' = RA\bar{R}. \tag{100}$$

Then, in equation (99), we have

$$\hat{R}(\partial_t + \nabla)\bar{R} = (R\partial\tilde{R})\hat{} = \widehat{\partial'} = \partial_{t'} + \nabla', \tag{101}$$

and, similarly,

$$\hat{R}(\phi - \mathbf{A})\bar{R} = (RA\tilde{R})\hat{} = \widehat{A'} = \phi' - \mathbf{A}'. \tag{102}$$

The transformed equation is, therefore, of the form

$$\hbar(\partial_{t'} + \nabla')\psi'\mathbf{ie}_3 = m\hat{\psi}' + e(\phi' - \mathbf{A}')\psi', \tag{103}$$

where the transformed spinor is

$$\psi' = R\psi. \tag{104}$$

The APS version of the Dirac equation has, as expected, Lorentz covariance.

4.3. Bilinear covariants

The electromagnetic and kinematic properties of the electron are described in Dirac theory in terms of quantities that are constructed as bilinear expressions involving the ψ spinor [26]. The so-called *bilinear covariant* quantities of Dirac theory are given, in terms of the APS, by [18, 19]

$$\sigma + \boldsymbol{i}\omega = \rho \mathrm{e}^{\mathrm{i}\beta} = \psi\bar{\psi}, \qquad j = \psi\tilde{\psi}, \qquad S = -\psi\boldsymbol{i}\mathbf{e}_3\bar{\psi}, \qquad s = \psi\mathbf{e}_3\tilde{\psi}, \tag{105}$$

where we define

$$\rho = \sqrt{\sigma^2 + \omega^2}, \qquad \tan\beta = \frac{\omega}{\sigma}. \tag{106}$$

The bilinear covariants can also be used in order to classify spinors, as in [11].

It is interesting to note that ψ can be written, using ρ and β defined above, as

$$\psi = \sqrt{\rho}\,\mathrm{e}^{\mathrm{i}\beta/2}R, \tag{107}$$

where R is an element of $\mathcal{C}\ell_3$ such that $R\bar{R} = \bar{R}R = 1$, that is, R describes a Lorentz transformation—see equation (60).

The so-called Fierz identities [28] can be easily proven using the APS. These identities, in terms of the APS, are

$$|j|^2 = \sigma^2 + \omega^2, \qquad |j|^2 = -|s|^2, \tag{108}$$

$$j \bullet s = 0, \qquad j \mathbin{\bar{\wedge}} s = -(\omega + \mathrm{i}\sigma)S, \tag{109}$$

where we use equations (54) and (57), and $|j|^2 = j\bar{j}$, $|s|^2 = s\bar{s}$. The proof of these identities can be easily done using equation (107). In fact, for the first one, we have

$$j\bar{j} = \psi\tilde{\psi}\hat{\psi}\bar{\psi} = \psi\widehat{(\bar{\psi}\psi)}\bar{\psi} = \psi\rho(\mathrm{e}^{\mathrm{i}\beta/2})\hat{}\,\bar{\psi} = \rho\mathrm{e}^{-\mathrm{i}\beta/2}\psi\bar{\psi} = \rho\mathrm{e}^{-\mathrm{i}\beta/2}\rho\mathrm{e}^{\mathrm{i}\beta/2} = \rho^2. \tag{110}$$

The second one follows from

$$s\bar{s} = \psi\mathbf{e}_3\tilde{\psi}\hat{\psi}(-\mathbf{e}_3)\bar{\psi} = -\rho\mathrm{e}^{-\mathrm{i}\beta/2}\psi(\mathbf{e}_3)^2\bar{\psi} = -\rho\mathrm{e}^{-\mathrm{i}\beta/2}\rho\mathrm{e}^{\mathrm{i}\beta/2} = -\rho^2. \tag{111}$$

The third and fourth ones follow from

$$j\bar{s} = \psi\tilde{\psi}\hat{\psi}(-\mathbf{e}_3)\bar{\psi} = -\rho\mathrm{e}^{-\mathrm{i}\beta/2}\psi\mathbf{e}_3\bar{\psi} = -\boldsymbol{i}(\sigma - \boldsymbol{i}\omega)S, \tag{112}$$

$$s\bar{j} = \psi\mathbf{e}_3\tilde{\psi}\hat{\psi}\bar{\psi} = \rho\mathrm{e}^{-\mathrm{i}\beta/2}\psi\bar{\psi} = \boldsymbol{i}(\sigma - \boldsymbol{i}\omega)S, \tag{113}$$

and the use of equations (54) and (57).

5. Conservation laws

The conservation laws in Dirac theory have been exploited using the spacetime algebra in [15, 29]. One interesting aspect in those approaches is that the conservation laws have been proved directly from the Dirac equation, and not from a Lagrangian one. In this section we follow the same approach and prove the conservation of probability and the conservation of energy–momentum directly from the Dirac equation, but using the APS.

5.1. Conservation of probability

A conservation law in the APS can be written in the form

$$\langle(\partial_t + \nabla)(j_0 + \boldsymbol{j})\rangle_0 = 0, \tag{114}$$

where $j_0 \in \bigwedge_0$ is a density and $\boldsymbol{j} \in \bigwedge_1$ a three-dimensional current, as can be easily verified; indeed, we have

$$\langle(\partial_t + \nabla)(j_0 + \boldsymbol{j})\rangle_0 = \langle\partial_t j_0 + \nabla j_0 + \partial_t \boldsymbol{j} + \nabla\cdot\boldsymbol{j} + \nabla\wedge\boldsymbol{j}\rangle_0 = \partial_t j_0 + \nabla\cdot\boldsymbol{j} = 0, \tag{115}$$

where we use equation (32) and the fact that $\partial_t \boldsymbol{j} \in \bigwedge_1$, $\nabla j_0 = \text{grad}\, j_0 \in \bigwedge_1$ and $\nabla\wedge\boldsymbol{j} \in \bigwedge_2$.

Let us write a paravector as $\psi_1 M\psi_2$. Since a paravector has to satisfy $\psi_1 M\psi_2 = (\psi_1 M\psi_2)^{\sim} = \tilde{\psi}_2\tilde{M}\tilde{\psi}_1$, we see that it has to be of the form $\psi M\tilde{\psi}$ with $M = \tilde{M}$. From equation (105), we have two 4-dimensional currents of the form $\psi M\tilde{\psi}$, for $M = \{1, \mathbf{e}_3\}$, in Dirac theory. Let us see if there is any conservation law associated with these quantities. In order to do this, we will look for an expression for $\langle(\partial_t + \nabla)(\psi M\tilde{\psi})\rangle$ using only the Dirac equation. For the sake of generality, we consider an arbitrary paravector M.

Let us apply grade involution to Dirac equation equation (69), that is,

$$\hbar(\partial_t - \nabla)\hat{\psi}\boldsymbol{i}\mathbf{e}_3 = m\psi + e(\phi + \mathbf{A})\hat{\psi}. \tag{116}$$

If we multiply it by $M\tilde{\psi}$ on the right, and then apply $(\partial_t + \nabla)$ to the result, we obtain

$$\begin{aligned} m(\partial_t + \nabla)(\psi M\tilde{\psi}) \;=\; & \hbar(\Box\hat{\psi})\boldsymbol{i}\mathbf{e}_3 M\tilde{\psi} + \hbar((\partial_t - \nabla)\hat{\psi})\boldsymbol{i}\mathbf{e}_3 M\partial_t\tilde{\psi} \\ & + \hbar\mathbf{e}^j((\partial_t - \nabla)\hat{\psi})\boldsymbol{i}\mathbf{e}_3 M\partial_j\tilde{\psi} - e[(\partial_t + \nabla)((\phi + \mathbf{A})\hat{\psi})]M\tilde{\psi} \\ & - e[(\phi + \mathbf{A})\hat{\psi}]M\partial_t\tilde{\psi} - e\mathbf{e}^j[(\phi + \mathbf{A})\hat{\psi}]M\partial_j\tilde{\psi}. \end{aligned} \tag{117}$$

But, from Dirac equation equation (69), if we apply the operator $(\partial_t - \nabla)$ on its left, we have

$$\Box\psi = -\frac{m^2}{\hbar^2}\psi - \frac{em}{\hbar^2}(\phi + \mathbf{A})\hat{\psi} - \frac{e}{\hbar}(\partial_t - \nabla)[(\phi - \mathbf{A})\psi]\boldsymbol{i}\mathbf{e}_3. \tag{118}$$

If we apply grade involution to equation (118), and use the result in equation (117), and after this use again Dirac equation equation (116), it follows that

$$(\partial_t + \nabla)(\psi M\tilde{\psi}) = -\frac{m}{\hbar}\hat{\psi}\boldsymbol{i}\mathbf{e}_3 M\tilde{\psi} - \frac{e}{\hbar}(\phi - \mathbf{A})\psi\boldsymbol{i}\mathbf{e}_3 M\tilde{\psi} + \psi M\partial_t\tilde{\psi} + \mathbf{e}^j\psi M\partial_j\tilde{\psi}. \tag{119}$$

Now we calculate the scalar part of this expression. First, we note that, because of equation (39), we can write

$$\langle\psi M\partial_t\tilde{\psi} + \mathbf{e}^j\psi M\partial_j\tilde{\psi}\rangle_0 = \langle\psi M[\tilde{\psi}\overleftarrow{(\partial_t + \nabla)}]\rangle_0, \tag{120}$$

where we denote

$$\tilde{\psi}\overleftarrow{(\partial_t + \nabla)} = \partial_t\tilde{\psi} + \partial_j\tilde{\psi}\mathbf{e}^j. \tag{121}$$

If we apply the reversion operation to the Dirac equation as in equation (69), we obtain

$$\hbar\tilde{\psi}\overleftarrow{(\partial_t + \nabla)} = \boldsymbol{i}\mathbf{e}_3 m\tilde{\psi} + e\boldsymbol{i}\mathbf{e}_3\tilde{\psi}(\phi - \mathbf{A}). \tag{122}$$

Now, if we take the scalar part of equation (119), and use equation (122) in equation (120), and use again equation (39), it follows that

$$\begin{aligned}\langle(\partial_t + \nabla)(\psi M\tilde{\psi})\rangle_0 &= -\frac{m}{\hbar}\langle \boldsymbol{i}\mathbf{e}_3 M\tilde{\psi}\hat{\psi}\rangle_0 + \frac{m}{\hbar}\langle \boldsymbol{i}M\mathbf{e}_3\bar{\psi}\psi\rangle_0 \\ &\quad + \frac{em}{\hbar}\langle(\phi - \mathbf{A})\psi\boldsymbol{i}[M, \mathbf{e}_3]\tilde{\psi}\rangle_0,\end{aligned} \tag{123}$$

where $[M, \mathbf{e}_3] = M\mathbf{e}_3 - \mathbf{e}_3 M$. We can simplify this expression if we remember equation (105), that is,

$$\bar{\psi}\psi = \sigma + \boldsymbol{i}\omega, \quad \tilde{\psi}\hat{\psi} = \widehat{\bar{\psi}\psi} = \sigma - \boldsymbol{i}\omega. \tag{124}$$

Using this in equations (123) and (39) again, we get that

$$\langle(\partial_t + \nabla)(\psi M\tilde{\psi})\rangle_0 = -\frac{2m\omega}{\hbar}\langle \mathbf{e}_3 M\rangle_0 + \frac{em}{\hbar}\langle(\phi - \mathbf{A})\psi\boldsymbol{i}[M, \mathbf{e}_3]\tilde{\psi}\rangle_0, \tag{125}$$

Let us consider $M = 1$. Since $\langle \mathbf{e}_3 1\rangle_0 = 0$ and $[1, \mathbf{e}_3] = 0$, it follows that

$$\langle(\partial_t + \nabla)\psi\tilde{\psi}\rangle_0 = 0, \tag{126}$$

that is, the probability current paravector $j = \psi\tilde{\psi}$ is conserved in Dirac theory. For $M = \mathbf{e}_3$, we have $\langle {\mathbf{e}_3}^2\rangle_0 = 1$ and $[\mathbf{e}_3, \mathbf{e}_3] = 0$, and then

$$\langle(\partial_t + \nabla)\psi\mathbf{e}_3\tilde{\psi}\rangle_0 = -\frac{2m\omega}{\hbar} = -\frac{2m\rho}{\hbar}\sin\beta, \tag{127}$$

that is, the spin density $\psi\mathbf{e}_3\tilde{\psi}$ is not conserved, unless $\beta = 0$ or $\beta = \pi$. It is interesting to observe that the free particle solutions of the Dirac equation satisfy $\beta = 0$ and $\beta = \pi$ [27]. There are two other interesting relations that follow from equation (125) for $M = \mathbf{e}_1$ and $M = \mathbf{e}_2$, namely

$$\langle(\partial_t + \nabla)\psi\mathbf{e}_1\tilde{\psi}\rangle_0 = \frac{2me}{\hbar}\langle(\phi - \mathbf{A})\psi\mathbf{e}_2\tilde{\psi}\rangle_0 \tag{128}$$

$$\langle(\partial_t + \nabla)\psi\mathbf{e}_2\tilde{\psi}\rangle_0 = -\frac{2me}{\hbar}\langle(\phi - \mathbf{A})\psi\mathbf{e}_1\tilde{\psi}\rangle_0. \tag{129}$$

5.2. Conservation of energy–momentum

Another important conservation law is the energy–momentum one. The above discussion shows that it cannot be associated with a paravector of the form $\psi M\tilde{\psi}$; it has to be of a different form. In order to obtain the expression for the conservation of energy–momentum, we first take into account the fact that, associated with the electric current paravector ej, there is the electromagnetic field $F = \mathbf{E} + \boldsymbol{i}\mathbf{B}$, satisfying Maxwell's equations, which in APS formalism is written as

$$(\partial_t + \nabla)(\mathbf{E} + \boldsymbol{i}\mathbf{B}) = e(\rho - \mathbf{j}). \tag{130}$$

In the appendix, we show that the scalar, vector, bivector and trivector parts of this single multivector equation correspond to the usual four Maxwell's equations. We also show, in the appendix, how to obtain from equation (130) the Poynting theorem in the form

$$\partial_t\mathcal{U} + \partial_j\mathcal{U}^j + f = 0, \tag{131}$$

where $\mathcal{U}$ is the paravector defined as $\mathcal{U} = u + \mathbf{g}$, with u and $\mathbf{g}$ being the energy and momentum densities, respectively, and $\mathcal{U}^j$ $(j = 1, 2, 3)$ are the paravectors defined as $\mathcal{U}^j = S^j - \tau^{jk}\mathbf{e}_k$ where S^j are the components of the energy flux vector $\mathbf{S}$ (the Poynting vector) and $\tau^{jk}\mathbf{e}_k$ is the flux vector associated with the g^j component of $\mathbf{g} = g^j\mathbf{e}_j$ (and $\mathbf{g} = \mathbf{S}$ since $c = 1$). Moreover, f is given by

$$f = e\langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\rho + \mathbf{j})\rangle_{0\oplus 1} = e\mathbf{E}\cdot\mathbf{j} + e(\rho\mathbf{E} + \mathbf{j}\times\mathbf{B}), \tag{132}$$

where we identify the vector part of f with the Lorentz force density.

Conservation of energy–momentum follows if we obtain from the Dirac equation an expression like

$$\partial_t\mathcal{T} + \partial_j\mathcal{T}^j = f, \tag{133}$$

where $\mathcal{T}$ and $\mathcal{T}^j$ ($j = 1, 2, 3$) are paravectors that play the role of mechanical counterparts of the quantities as in equation (131), in such a way that

$$\partial_t(\mathcal{U} + \mathcal{T}) + \partial_j(\mathcal{U}^j + \mathcal{T}^j) = 0. \tag{134}$$

An expression like equation (133) can be obtained from equation (118). In order to do it, we first note that the last term on the RHS of equation (118) can be written as

$$\begin{aligned}(\partial_t - \nabla)[(\phi - \mathbf{A})\psi] = {} & [(\partial_t - \nabla)(\phi - \mathbf{A})]\psi - (\phi + \mathbf{A})[(\partial_t + \nabla)\psi] \\ & + 2(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi.\end{aligned} \tag{135}$$

If we use the Dirac equation in the second term of the RHS of equation (135), and then use the result in equation (118), we obtain

$$\begin{aligned}\Box\psi \; = \; & -\frac{m^2}{\hbar^2}\psi - \frac{e}{h}[(\partial_t - \nabla)(\phi - \mathbf{A})]\psi\boldsymbol{i}\mathbf{e}_3 + \frac{e^2}{\hbar^2}(\phi^2 - \mathbf{A}^2)\psi \\ & - 2\frac{e}{\hbar}(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi\boldsymbol{i}\mathbf{e}_3.\end{aligned} \tag{136}$$

Now, let us multiply equation (136) on the right side by $H\psi^\dagger$, where $\psi^\dagger$ denotes either $\tilde{\psi}$ or $\bar{\psi}$, and H is an element of $\bigwedge_2 \oplus \bigwedge_3$ such that $[H, \mathbf{e}_3]$. After some calculations, which we leave for an appendix, we obtain

$$\partial_t\langle\hbar\partial_t\psi H\tilde{\psi} + e\phi\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus 1} - \partial_j\langle\hbar\partial^j\psi H\tilde{\psi} - eA^j\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus 1} = -e\langle F\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus 1}. \tag{137}$$

We want to identify the RHS of equation (137) with f as in equation (132). Since $\psi\tilde{\psi} = \rho + \mathbf{j}$, we need to choose $H = \mathbf{i}\mathbf{e}_3$ for this to hold. Then, we have

$$\partial_t\langle(\hbar\partial_t\psi\boldsymbol{i}\mathbf{e}_3 - e\phi\psi)\tilde{\psi}\rangle_{0\oplus 1} + \partial_j\langle(-\hbar\partial^j\psi\boldsymbol{i}\mathbf{e}_3 - eA^j\psi)\tilde{\psi}\rangle_{0\oplus 1} = f. \tag{138}$$

Note that this equation has the form of equation (133). Let us introduce the following notation:

$$T^{00} = \langle(\hbar\partial_t\psi\boldsymbol{i}\mathbf{e}_3 - e\phi\psi)\tilde{\psi}\rangle_0, \qquad T^{0j} = \langle(\hbar\partial_t\psi\boldsymbol{i}\mathbf{e}_3 - e\phi\psi)\tilde{\psi}\mathbf{e}^j\rangle_0, \tag{139}$$

$$T^{j0} = \langle(-\hbar\partial^j\psi\boldsymbol{i}\mathbf{e}_3 - eA^j\psi)\tilde{\psi}\rangle_0, \qquad T^{jk} = \langle(-\hbar\partial^j\psi\boldsymbol{i}\mathbf{e}_3 - eA^j\psi)\tilde{\psi}\mathbf{e}^k\rangle_0. \tag{140}$$

Now we can define the paravectors $\mathcal{T}$ and $\mathcal{T}^j$ ($j = 1, 2, 3$) as

$$\mathcal{T} = T^{00} + T^{0j}\mathbf{e}_j, \tag{141}$$

$$\mathcal{T}^j = T^{j0} + T^{jk}\mathbf{e}_k, \tag{142}$$

in such a way that equation (138) can be written as

$$\partial_t\mathcal{T} + \partial_j\mathcal{T}^j = f, \tag{143}$$

which is of the form of equation (133). Then, from equation (134) follows the conservation of the total (mechanical and electromagnetic) energy–momentum.

6. CPT and APS

Let us consider, in the formalism of the APS, the effect of the parity (P), time reversal (T) and charge conjugation (C) operations on the Dirac equation, as given by equation (69), and on the ψ spinor.

6.1. Parity

Under parity (P) transformation, the transformed Dirac equation must have the following form:

$$\hbar(\partial_t - \nabla)\mathsf{P}[\psi]\boldsymbol{i}\mathbf{e}_3 = m\widehat{\mathsf{P}[\psi]} + e(\phi + \mathbf{A})\mathsf{P}[\psi], \tag{144}$$

where $\mathsf{P}[\psi]$ is the parity-transformed ψ field. Now we need to find a transformation that takes the above equation in the original Dirac equation. Since there does *not* exist $U \in \mathcal{C}\ell_3$ such that $U\mathbf{e}_i = -\mathbf{e}_i U$ for $i = 1, 2, 3$, this transformation cannot be simply of the form $\mathsf{P}[\psi] = U\psi V$. However, since $\hat{\mathbf{e}}_i = -\mathbf{e}_i$, we can use grade involution as part of this transformation, and look for an expression for P in the form

$$\mathsf{P}[\psi] = U\hat{\psi}V. \tag{145}$$

where U and V are elements of $\mathcal{C}\ell_3$.

Using equation (145) in equation (144) and comparing the result with equation (69), we can see that we must have:

$$\text{(i) } U^{-1}\mathbf{e}_i\hat{U} = \mathbf{e}_i, \qquad \text{(ii) } U^{-1}\hat{U} = 1, \qquad \text{(iii) } \hat{V}\mathbf{e}_3 V^{-1} = \mathbf{e}_3, \qquad \text{(iv) } \hat{V}V^{-1} = 1, \tag{146}$$

where $i = 1, 2, 3$. Condition (i) implies that U belongs to the center of $\mathcal{C}\ell_3$, that is, $U \in \bigwedge_0 \oplus \bigwedge_3$. However, condition (ii) implies that $U = \hat{U}$, that is, U is an even element. Therefore, from these two conditions we conclude that $U = 1$. On the other hand, from condition (iii) we have four possibilities, $V = \{1, \mathbf{e}_3, \mathrm{i}\mathbf{e}_3, \mathrm{i}\}$. But condition (iv) implies that $\hat{V} = V$, an even element. Therefore, we have two possibilities: $V = \{1, \mathrm{i}\mathbf{e}_3\}$. In principle the most general expression for $\mathsf{P}(\psi)$ is a combination of these two cases, and can be written in the form

$$\mathsf{P}[\psi] = \hat{\psi}\mathrm{e}^{\alpha\mathrm{i}\mathbf{e}_3}, \tag{147}$$

where α is an arbitrary real constant, and we remember that $\mathrm{i}\mathbf{e}_3 = \mathbf{e}_{12}$.

Note that

$$\mathsf{P}^2[\psi] = \psi\ \mathrm{e}^{2\alpha\mathrm{i}\mathbf{e}_3}. \tag{148}$$

which is a solution of the original Dirac equation since the original ψ field appears to be multiplied by an arbitrary phase.

Remark. We have excluded the possibility that $V = \{\mathbf{e}_3, \boldsymbol{i}\}$ because of condition (iv) which demands V to be an even element. If, however, we have considered instead the alternative condition (iv-a) $\hat{V}V^{-1} = -1$ then V would be an odd element, and the solution for the parity transformation would be of the form

$$\mathfrak{P}[\psi] = \hat{\psi}\mathbf{e}_3\mathrm{e}^{\alpha\mathrm{i}\mathbf{e}_3}. \tag{149}$$

However, the condition (iv-a) $\hat{V}V^{-1} = -1$ would change the sign of the term with the electric coupling, that is, it would change $e \mapsto -e$. But this is not what is expected for a parity transformation! Nevertheless, a parity transformation of the form $\mathfrak{P}$ can be considered in the case of an *uncharged* particle.

6.2. Time reversal

Under time reversal (T), the transformed Dirac equation has the form

$$\hbar(-\partial_t + \nabla)\mathsf{T}[\psi]\boldsymbol{i}\mathbf{e}_3 = m\widehat{\mathsf{T}[\psi]} + e(\phi + \mathbf{A})\mathsf{T}[\psi]. \tag{150}$$

If we suppose that $\mathsf{T}[\psi] = U\hat{\psi}V$, we must have

$$\text{(i) } U^{-1}\mathbf{e}_i\hat{U} = \mathbf{e}_i, \qquad \text{(ii) } U^{-1}\hat{U} = 1, \qquad \text{(iii) } \hat{V}\mathbf{e}_3V^{-1} = -\mathbf{e}_3, \qquad \text{(iv)} \hat{V}V^{-1} = 1, \tag{151}$$

where $i = 1, 2, 3$. From conditions (i) and (ii) we conclude that $U = 1$. From condition (ii) we conclude that $V = \{\mathbf{e}_1, \mathbf{e}_2, \boldsymbol{i}\mathbf{e}_1, \boldsymbol{i}\mathbf{e}_2\}$, and from condition (iv) that V must be an even element.

The most general expression for $\mathsf{T}[\psi]$ is therefore

$$\mathsf{T}[\psi] = \hat{\psi}\boldsymbol{i}\mathbf{e}_1\mathrm{e}^{\beta\mathbf{i}\mathbf{e}_3}, \tag{152}$$

where β is an arbitrary real constant. It is interesting to note that

$$\mathsf{T}^2 = -1 \tag{153}$$

for all values of β in equation (152).

6.3. Charge conjugation

Under charge conjugation (C), the transformed Dirac equation has the form

$$\hbar(\partial_t + \nabla)\mathsf{C}[\psi]\boldsymbol{i}\mathbf{e}_3 = m\widehat{\mathsf{C}[\psi]} - e(\phi - \mathbf{A})\mathsf{C}[\psi]. \tag{154}$$

If we suppose that $\mathsf{C}[\psi] = U\psi V$, we have

$$\text{(i) } U^{-1}\mathbf{e}_iU = \mathbf{e}_i, \qquad \text{(ii) } U^{-1}\hat{U} = 1, \qquad \text{(iii) } V\mathbf{e}_3V^{-1} = -\mathbf{e}_3, \qquad \text{(iv) } \hat{V}V^{-1} = -1, \tag{155}$$

where $i = 1, 2, 3$, and then $U = 1$ and $V = \{\mathbf{e}_1, \mathbf{e}_2\}$. The most general expression for $\mathsf{C}[\psi]$ in this case is therefore

$$\mathsf{C}[\psi] = \psi\mathbf{e}_1\mathrm{e}^{\theta\mathbf{i}\mathbf{e}_3}. \tag{156}$$

We also note that

$$\mathsf{C}^2 = 1 \tag{157}$$

for all values of θ in equation (156).

6.4. Commutation relations and CPT

The commutation or anticommutation relations between P, T and C, as given by equations (147), (152) and (156), depend on the values of α, β and θ. Therefore, we can remove at least some of this arbitrariness, by requiring some definite commutation or anticommutation relations between P, T and C. For example, we can see that

$$[\mathsf{P}, \mathsf{T}](\psi) = 2\boldsymbol{i}\psi\mathbf{e}_2\mathrm{e}^{\beta\mathbf{i}\mathbf{e}_3}\sin\alpha,$$

and if $\sin\alpha = 0$, we have $[\mathsf{P}, \mathsf{T}] = 0$. Similar relations can be obtained in other cases.

Let us summarize the possibilities that we have:

$$\textbf{(i)} \quad [\mathsf{P}, \mathsf{T}] = 0 \Longleftrightarrow \sin\alpha = 0, \tag{158}$$

$$\textbf{(ii)}\quad \{\mathsf{P},\,\mathsf{T}\} = 0 \Longleftrightarrow \cos\alpha = 0, \tag{159}$$

$$\textbf{(iii)}\quad [\mathsf{P},\,\mathsf{C}] = 0 \Longleftrightarrow \cos\alpha = 0, \tag{160}$$

$$\textbf{(iv)}\quad \{\mathsf{P},\,\mathsf{C}\} = 0 \Longleftrightarrow \sin\alpha = 0, \tag{161}$$

$$\textbf{(v)}\quad [\mathsf{C},\,\mathsf{T}] = 0 \Longleftrightarrow \cos(\theta - \beta) = 0, \tag{162}$$

$$\textbf{(vi)}\quad \{\mathsf{C},\,\mathsf{T}\} = 0 \Longleftrightarrow \sin(\theta - \beta) = 0. \tag{163}$$

In relation to the composition of C, P and T, we have

$$\mathsf{CPT}(\psi) = \mathrm{i}\psi \mathrm{e}^{(\theta-\alpha-\beta)\mathbf{ie}_3}, \tag{164}$$

and therefore

$$\textbf{(i)}\quad (\mathsf{CPT})^2 = -1 \Longleftrightarrow \cos 2(\theta - \alpha - \beta) = 1, \tag{165}$$

$$\textbf{(ii)}\quad (\mathsf{CPT})^2 = 1 \Longleftrightarrow \cos 2(\theta - \alpha - \beta) = -1. \tag{166}$$

6.5. Four different sets for C, P and T

From the above subsection we can conclude that there are four different possible sets of definitions for the operators P, T and C, apart from the arbitrariness of a real constant β. These four sets are as follows:

(I) The operators in this case satisfy

$$[\mathsf{P},\,\mathsf{T}] = \{\mathsf{P},\,\mathsf{C}\} = \{\mathsf{C},\,\mathsf{T}\} = 0, \qquad \mathsf{P}^2 = 1,\ \mathsf{T}^2 = -1,\ \mathsf{C}^2 = 1,$$
$$(\mathsf{CPT})^2 = -1, \tag{167}$$

and they are defined as

$$\mathsf{P}[\psi] = \hat{\psi}, \qquad \mathsf{T}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}, \qquad \mathsf{C}[\psi] = \psi\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}. \tag{168}$$

(II) In this case the operators satisfy

$$\{\mathsf{P},\,\mathsf{T}\} = [\mathsf{P},\,\mathsf{C}] = [\mathsf{C},\,\mathsf{T}] = 0, \quad \mathsf{P}^2 = -1,$$
$$\mathsf{T}^2 = -1,\ \mathsf{C}^2 = 1, \quad (\mathsf{CPT})^2 = -1, \tag{169}$$

and they are defined as

$$\mathsf{P}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_3, \qquad \mathsf{T}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}, \qquad \mathsf{C}[\psi] = \psi\mathbf{e}_2\,\mathrm{e}^{\beta\mathbf{ie}_3}. \tag{170}$$

(III) In this case the operators satisfy

$$[\mathsf{P},\,\mathsf{T}] = \{\mathsf{P},\,\mathsf{C}\} = [\mathsf{C},\,\mathsf{T}] = 0, \qquad \mathsf{P}^2 = 1,\ \mathsf{T}^2 = -1,\ \mathsf{C}^2 = 1, \qquad (\mathsf{CPT})^2 = 1, \tag{171}$$

and they are defined as

$$\mathsf{P}[\psi] = \hat{\psi}, \qquad \mathsf{T}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}, \qquad \mathsf{C}[\psi] = \psi\mathbf{e}_2\,\mathrm{e}^{\beta\mathbf{ie}_3}. \tag{172}$$

(IV) In this case we have

$$\{\mathsf{P},\,\mathsf{T}\} = [\mathsf{P},\,\mathsf{C}] = \{\mathsf{C},\,\mathsf{T}\} = 0, \qquad \mathsf{P}^2 = -1,\ \mathsf{T}^2 = -1,\ \mathsf{C}^2 = 1,$$
$$(\mathsf{CPT})^2 = 1, \tag{173}$$

and they are defined as

$$\mathsf{P}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_3, \qquad \mathsf{T}[\psi] = \boldsymbol{i}\hat{\psi}\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}, \qquad \mathsf{C}[\psi] = \psi\mathbf{e}_1\,\mathrm{e}^{\beta\mathbf{ie}_3}. \tag{174}$$

Remark. Case (I) is the set of C, P and T operators used in Dirac theory [11] written in terms of the APS. Case (II) has the same relations, equation (169), that are satisfied by ELKO spinors [30–33]. ELKO spinors have been considered as candidates to describe dark matter and to probe non-trivial topologies in spacetime [34]. However, ELKO spinors are not Dirac spinors and do not satisfy the Dirac equation [35], and the fact that we have obtained the same commutation/anticommutation relations deserves additional studies. In order to have a better understanding of cases (III) and (IV), it is convenient to note that, in all the above cases, the operators CP have the same form, that is, all can be written as

$$\mathsf{CP}[\psi] = \hat{\psi}\mathbf{e}_1 e^{\delta i \mathbf{e}_3}, \tag{175}$$

where $\delta = \beta$ for cases (I) and (II), $\delta = \beta + \pi/2$ for case (III), and $\delta = \beta - \pi/2$ for case (IV). We also notice that, for cases (I) and (II), we have $\{\mathsf{CP}, \mathsf{T}\} = 0$, and for cases (III) and (IV), we have $[\mathsf{CP}, \mathsf{T}] = 0$. Moreover, in all cases, $(\mathsf{CP})^2 = -1$. Then, if we consider the operators CP, T and CPT, we can construct a multiplication table for cases (I), (II), (III) and (IV). According to Wigner [36, 37], there are four different classes of C, P and T operators, which can be summarized in the following table:[1]

	1	CP	T	CPT
1	1	CP	T	CPT
CP	CP	-1	CPT	$-$T
T	T	$\epsilon\epsilon'$CPT	$-\epsilon'$	$-\epsilon$CP
CPT	CPT	$-\epsilon\epsilon'$T	$-\epsilon'$CP	ϵ

The four Wigner classes correspond to $(\epsilon, \epsilon') = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$. The fact is that all the cases (I), (II), (III) and (IV) correspond to the Wigner classes with $\epsilon' = 1$; furthermore, cases (I) and (II) have $\epsilon = -1$ and cases (III) and (IV) have $\epsilon = 1$. The differences between cases (I) and (II) in one Wigner class and between cases (III) and (IV) in the other Wigner class can be described by the different definitions of the parity operator; that is, in cases (I) and (III) we have the standard definition $\mathsf{P}[\psi] = \hat{\psi}$, while in cases (II) and (IV) we have the nonstandard definition $\mathsf{P}[\psi] = i\hat{\psi}\mathbf{e}_3$. Possible physical interpretations of cases (III) and (IV) require further investigations; we must bear in mind, however, the possibility that they may be just spurious solutions, due to the fact that we are working with multivectors. Anyway, we believe we are facing an interesting scenario to compare different approaches to Dirac theory.

6.5.1. Matrix representations of C, P and T. Let us use equations (93) and (95) in order to write the matrix representation of the C, P and T operators acting on Weyl spinors. We will use, for the action of an operator O, the notation

$$\mathsf{O}(\psi) = \big((\boldsymbol{\psi}^O)_{\mathrm{L}} \quad - i\sigma_2(\boldsymbol{\psi}^O)^*_{\mathrm{R}}\big). \tag{176}$$

(I) The matrix representation of equation (168) is

$$(\boldsymbol{\psi}^P)_{\mathrm{L}} = \boldsymbol{\psi}_{\mathrm{R}}, \quad (\boldsymbol{\psi}^T)_{\mathrm{L}} = e^{i\beta}\sigma_2\boldsymbol{\psi}^*_{\mathrm{L}}, \quad (\boldsymbol{\psi}^C)_{\mathrm{L}} = -e^{i\beta} i\sigma_2\boldsymbol{\psi}^*_{\mathrm{R}}, \tag{177}$$

$$(\boldsymbol{\psi}^P)_{\mathrm{R}} = \boldsymbol{\psi}_{\mathrm{L}}, \quad (\boldsymbol{\psi}^T)_{\mathrm{R}} = e^{i\beta}\sigma_2\boldsymbol{\psi}^*_{\mathrm{R}}, \quad (\boldsymbol{\psi}^C)_{\mathrm{R}} = e^{i\beta} i\sigma_2\boldsymbol{\psi}^*_{\mathrm{L}}. \tag{178}$$

[1] Our notation differs slightly from [37]; the correspondence is $\mathsf{CP} = i\Sigma$ and $\mathsf{T} = -i\theta$, where Σ and θ are the notations used in [37].

(II) The matrix representation of equation (170) is

$$(\psi^P)_\mathrm{L} = \mathrm{i}\psi_\mathrm{R}, \quad (\psi^T)_\mathrm{L} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{L}^*, \quad (\psi^C)_\mathrm{L} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{R}^*, \tag{179}$$

$$(\psi^P)_\mathrm{R} = -\mathrm{i}\psi_\mathrm{L}, \quad (\psi^T)_\mathrm{R} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{R}^*, \quad (\psi^C)_\mathrm{R} = -\mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{L}^*. \tag{180}$$

(III) The matrix representation of equation (172) is

$$(\psi^P)_\mathrm{L} = \psi_\mathrm{R}, \quad (\psi^T)_\mathrm{L} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{L}^*, \quad (\psi^C)_\mathrm{L} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{R}^*, \tag{181}$$

$$(\psi^P)_\mathrm{R} = \psi_\mathrm{L}, \quad (\psi^T)_\mathrm{R} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{R}^*, \quad (\psi^C)_\mathrm{R} = -\mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{L}^*. \tag{182}$$

(IV) The matrix representation of equation (174) is

$$(\psi^P)_\mathrm{L} = \mathrm{i}\psi_\mathrm{R}, \quad (\psi^T)_\mathrm{L} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{L}^*, \quad (\psi^C)_\mathrm{L} = -\mathrm{e}^{\mathrm{i}\beta}\mathrm{i}\sigma_2\psi_\mathrm{R}^*, \tag{183}$$

$$(\psi^P)_\mathrm{R} = -\mathrm{i}\psi_\mathrm{L}, \quad (\psi^T)_\mathrm{R} = \mathrm{e}^{\mathrm{i}\beta}\sigma_2\psi_\mathrm{R}^*, \quad (\psi^C)_\mathrm{R} = \mathrm{e}^{\mathrm{i}\beta}\mathrm{i}\sigma_2\psi_\mathrm{L}^*. \tag{184}$$

7. Conclusions

The assertion that the Dirac equation cannot be written solely using the algebra of Pauli matrices is not completely true. It is true only if we suppose that the quantity $\boldsymbol{\beta}(\psi)$ in the mass term is given by a matrix product. We have shown that there is another possibility given by $\boldsymbol{\beta}(\psi) = \mathrm{adj}(\psi^\dagger)\sigma_3$, where $\mathrm{adj}(\psi^\dagger)$ is the classical adjoint (adjugate) matrix of $\psi^\dagger$ and $\dagger$ denotes Hermitian conjugation. We have interpreted this other possibility in terms of the APS, and used this fact to write the Dirac equation in the APS, and exploited some of its properties. The Dirac spinor, in this formalism, is an arbitrary multivector of $\mathcal{C}\ell_3$. Left and right spinors appear naturally as the left and right columns of the matrix representation of the Dirac spinor. The proof of the Fierz identities is a remarkable example of the fact that the possibility of having an invertible element (the matrix ψ) facilitates calculations in the APS formalism. The conservation of probability and of energy–momentum followed from the Dirac equation, without the need for using the Lagrangian formalism. In the discussion of energy–momentum conservation, we touch on the use of the APS in electromagnetism (discussed in more detail in the appendix). We have also discussed how to implement the parity, time reversal and charge conjugation operations in the APS formalism, and arrived at four different sets of P, T and C operators, which, to the best of our knowledge, have not appeared before in the context of Dirac theory.

Many other aspects of Dirac theory can be discussed using the APS formalism, but we believe that, with this presentation, we have made our point, namely that the use of the APS formalism has some advantages over the usual formalism of Dirac matrices and column spinors, and should be considered as an alternative and powerful tool in relativistic quantum mechanics, and that could have been in current use if at some point in the development of the theory of the spin, some choices, in particular those involving the use of quaternions, had prevailed.

Appendix A. Proof of equation (137)

In this appendix, in order to prove equation (137), we use an approach similar to [15], adapted to the APS [38]. Let us start with equation (136), that is,

$$\begin{aligned}\Box\psi &= -\frac{m^2}{\hbar^2}\psi - \frac{e}{h}[(\partial_t - \nabla)(\phi - \mathbf{A})]\psi \boldsymbol{i}\mathbf{e}_3 + \frac{e^2}{\hbar^2}(\phi^2 - \mathbf{A}^2)\psi \\ &\quad - 2\frac{e}{\hbar}(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi\boldsymbol{i}\mathbf{e}_3.\end{aligned} \tag{185}$$

which we will multiply on the right by $H\psi^\dagger$, where $\psi^\dagger$ denotes either $\tilde{\psi}$ or $\bar{\psi}$, and H, for the moment, is an arbitrary element of the APS. Let us consider the expression of the form $\Box\psi H\psi^\dagger$. We want to consider its paravector part, which we will write as $\langle\Box\psi H\psi^\dagger\rangle_{0\oplus1}$. Since a paravector a satisfies $a = \tilde{a}$, we have

$$\begin{aligned}\langle\Box\psi H\psi^\dagger\rangle_{0\oplus1} &= \frac{1}{2}\langle\Box\psi H\psi^\dagger + \tilde{\psi}^\dagger\tilde{H}\Box\tilde{\psi}\rangle_{0\oplus1} \\ &= \frac{1}{2}[\partial_t\langle\partial_t\psi H\psi^\dagger + \tilde{\psi}^\dagger\tilde{H}\partial_t\tilde{\psi}\rangle_{0\oplus1} - \partial_j\langle\partial^j\psi H\psi^\dagger + \tilde{\psi}^\dagger\tilde{H}\partial^j\tilde{\psi}\rangle_{0\oplus1} \\ &\quad - \langle\partial_t\psi H\partial_t\psi^\dagger + \partial_t\tilde{\psi}^\dagger\tilde{H}\partial_t\tilde{\psi}\rangle_{0\oplus1} + \langle\partial^j\psi H\partial_j\psi^\dagger + \partial_j\tilde{\psi}\tilde{H}\partial^j\tilde{\psi}\rangle_{0\oplus1}],\end{aligned} \tag{186}$$

where, because of the use of the summation convention, we have written $\nabla^2\psi$ as $\nabla^2\psi = \partial_j\partial^j\psi$ with $\partial^j\psi = \partial_j\psi$. In order to have an expression of the form of the RHS of equation (133), we must have

$$\langle\partial_t\psi H\partial_t\psi^\dagger + \partial_t\tilde{\psi}^\dagger\tilde{H}\partial_t\tilde{\psi}\rangle_{0\oplus1} - \langle\partial^j\psi H\partial_j\psi^\dagger + \partial_j\tilde{\psi}\tilde{H}\partial^j\tilde{\psi}\rangle_{0\oplus1} = 0, \tag{187}$$

which is satisfied if

$$\psi^\dagger = \tilde{\psi}, \qquad \tilde{H} = -H. \tag{188}$$

The condition $\tilde{H} = -H$ means that H is an element of $\bigwedge_2 \oplus \bigwedge_3$. Then, in this case, it follows that

$$\langle\Box\psi H\tilde{\psi}\rangle_{0\oplus1} = \partial_t\langle\partial_t\psi H\tilde{\psi}\rangle_{0\oplus1} - \partial_j\langle\partial^j\psi H\tilde{\psi}\rangle_{0\oplus1}, \tag{189}$$

where we have used the fact that paravectors satisfy $a = \tilde{a}$. Because of this, we also have

$$\langle\psi H\tilde{\psi}\rangle_{0\oplus1} = \widetilde{\langle\psi H\tilde{\psi}\rangle}_{0\oplus1} = -\langle\psi H\tilde{\psi}\rangle_{0\oplus1} \implies \langle\psi H\tilde{\psi}\rangle_{0\oplus1} = 0. \tag{190}$$

Now, if we take the paravector part of the result of equation (185) multiplied on the right by $H\tilde{\psi}$, and use equations (189) and (190), we obtain

$$\begin{aligned}\partial_t\langle\partial_t\psi H\tilde{\psi}\rangle_{0\oplus1} - \partial_j\langle\partial^j\psi H\tilde{\psi}\rangle_{0\oplus1} &= -\frac{e}{\hbar}\langle[(\partial_t - \nabla)(\phi - \mathbf{A})]\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} \\ &\quad - 2\frac{e}{\hbar}\langle[(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi]\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}.\end{aligned} \tag{191}$$

The first term of the RHS can be written as

$$\langle[(\partial_t - \nabla)(\phi - \mathbf{A})]\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} = \langle(\partial_t\phi + \nabla\cdot\mathbf{A} + \mathbf{E} + \boldsymbol{i}\mathbf{B})\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}, \tag{192}$$

where we use

$$\mathbf{E} = -\partial_t\mathbf{A} - \nabla\phi, \qquad \boldsymbol{i}\mathbf{B} = \nabla\wedge\mathbf{A}. \tag{193}$$

In relation to the second term on the RHS, we can write

$$\begin{aligned}\langle[(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi]\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} &= \frac{1}{2}\langle[(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi]\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi} \\ &\quad + \psi\boldsymbol{i}H\mathbf{e}_3[\tilde{\psi}\overleftarrow{(\phi\partial_t + \mathbf{A}\cdot\nabla)}]\rangle_{0\oplus1} \\ &= \frac{1}{2}\phi\langle\partial_t\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} + \frac{1}{2}A^j\langle\partial_j\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} \\ &\quad + \frac{1}{2}\phi\langle\psi\boldsymbol{i}H\mathbf{e}_3\partial_t\tilde{\psi}\rangle_{0\oplus1} + \frac{1}{2}A^j\langle\psi\boldsymbol{i}H\mathbf{e}_3\partial_j\tilde{\psi}\rangle_{0\oplus1}.\end{aligned} \tag{194}$$

If we suppose that

$$[H, \mathbf{e}_3] = 0, \tag{195}$$

it follows that

$$\langle[(\phi\partial_t + \mathbf{A}\cdot\nabla)\psi]\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} = \frac{1}{2}\phi\partial_t\langle\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} + \frac{1}{2}(\mathbf{A}\cdot\nabla)\langle\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}. \tag{196}$$

If we use equations (192) and (196) in equation (191), and use the fact that

$$\langle\partial_t\phi\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} + \phi\partial_t\langle\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} = \partial_t\langle\phi\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}, \tag{197}$$

$$(\nabla\cdot\mathbf{A})\langle\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} + (\mathbf{A}\cdot\nabla)\langle\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} = \partial_j\langle A^j\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}, \tag{198}$$

we obtain

$$\partial_t\langle\hbar\partial_t\psi H\tilde{\psi} + e\phi\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} - \partial_j\langle\hbar\partial^j\psi H\tilde{\psi} - eA^j\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1} = -e\langle F\psi\boldsymbol{i}\mathbf{e}_3 H\tilde{\psi}\rangle_{0\oplus1}. \tag{199}$$

with $F = \mathbf{E} + \boldsymbol{i}\mathbf{B}$.

Appendix B. Electromagnetism in the APS

Maxwell's equations can be written as a single equation in the APS, namely

$$(\partial_t + \nabla)(\mathbf{E} + \boldsymbol{i}\mathbf{B}) = \rho_0 - \mathbf{j}_0, \tag{200}$$

where ρ_0 is the electric charge density and $\mathbf{j}_0$ is the electric charge current. The usual set of four equations follows from the projection of this multivector equation in its k-vector parts for $k = 0, 1, 2, 3$. If we expand the LHS of equation (200), we have

$$\partial_t\mathbf{E} + \mathrm{i}\partial_t\mathbf{B} + \operatorname{div}\mathbf{E} + \mathrm{i}\operatorname{rot}\mathbf{E} + \mathrm{i}\operatorname{div}\mathbf{B} + \mathrm{i}^2\operatorname{rot}\mathbf{B} = \rho_0 - \mathbf{j}_0. \tag{201}$$

where we have used equation (42), and taking its k-vector parts we obtain the usual Maxwell's equations,

$$\begin{aligned}\langle\ \rangle_0 &\Rightarrow \operatorname{div}\mathbf{E} = \rho_0,\\ \langle\ \rangle_1 &\Rightarrow \operatorname{rot}\mathbf{B} - \partial_t\mathbf{E} = \mathbf{j}_0,\\ \langle\ \rangle_2 &\Rightarrow \boldsymbol{i}\operatorname{rot}\mathbf{E} + \boldsymbol{i}\partial_t\mathbf{B} = 0,\\ \langle\ \rangle_3 &\Rightarrow \boldsymbol{i}\operatorname{div}\mathbf{B} = 0.\end{aligned} \tag{202}$$

The Poynting theorem follows easily from Maxwell's equations in the form of equation (200). First, for convenience, let us apply grade involution to equation (200), writing Maxwell's equations in an equivalent form as

$$(\partial_t - \nabla)(-\mathbf{E} + \boldsymbol{i}\mathbf{B}) = \rho_0 + \mathbf{j}_0, \tag{203}$$

Now, let us multiply it on the left by $\mathbf{E} + \boldsymbol{i}\mathbf{B}$ and take the paravector part,

$$\langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\partial_t - \nabla)(-\mathbf{E} + \boldsymbol{i}\mathbf{B})\rangle_{0\oplus 1} = \langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\rho_0 + \mathbf{j}_0\rangle_{0\oplus 1}. \tag{204}$$

The content of this expression is known as the Poynting theorem. Let us prove this. The RHS is

$$\langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})\rho_0 + \mathbf{j}_0\rangle_{0\oplus 1} = \mathbf{E}\cdot\mathbf{j}_0 + \rho_0\mathbf{E} + \mathbf{j}_0 \times \mathbf{B} = f. \tag{205}$$

In order to calculate the LHS, we will use $\langle A\rangle_{0\oplus 1} = (A + \tilde{A})/2$, that is

$$\begin{aligned}\langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\partial_t - \nabla)(-\mathbf{E} + \boldsymbol{i}\mathbf{B})\rangle_{0\oplus 1} &= \frac{1}{2}[(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\partial_t - \nabla)(-\mathbf{E} + \boldsymbol{i}\mathbf{B}) \\ &\quad + (-\mathbf{E} - \boldsymbol{i}\mathbf{B})\overleftarrow{(\partial_t - \nabla)}(\mathbf{E} - \boldsymbol{i}\mathbf{B})].\end{aligned} \tag{206}$$

After expanding the terms, we obtain

$$\begin{aligned}\langle(\mathbf{E} + \boldsymbol{i}\mathbf{B})(\partial_t - \nabla)(-\mathbf{E} + \boldsymbol{i}\mathbf{B})\rangle_{0\oplus 1} &= -\partial_t\left(\frac{\mathbf{E}^2 + \mathbf{B}^2}{2} + \frac{\boldsymbol{i}\mathbf{B}\mathbf{E} - \boldsymbol{i}\mathbf{E}\mathbf{B}}{2}\right) \\ &\quad + \partial_j\left(\frac{\boldsymbol{i}\mathbf{B}\mathbf{e}^j\mathbf{E} - \boldsymbol{i}\mathbf{E}\mathbf{e}^j\mathbf{B}}{2} + \frac{\mathbf{E}\mathbf{e}^j\mathbf{E} + \mathbf{B}\mathbf{e}^i\mathbf{B}}{2}\right).\end{aligned} \tag{207}$$

The term

$$u = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) \tag{208}$$

is the energy density of the electromagnetic fields, and the term

$$\frac{\boldsymbol{i}\mathbf{B}\mathbf{E} - \boldsymbol{i}\mathbf{E}\mathbf{B}}{2} = -\boldsymbol{i}\mathbf{E}\wedge\mathbf{B} = \mathbf{E}\times\mathbf{B} = \mathbf{S} = \mathbf{g}, \tag{209}$$

represents the momentum density $\mathbf{g}$ of the electromagnetic field, which is equal to the Poynting vector $\mathbf{S}$ since we are using $c = 1$. The next term in equation (207) is

$$\begin{aligned}\frac{\boldsymbol{i}\mathbf{B}\mathbf{e}^j\mathbf{E} - \mathrm{i}\mathbf{E}\mathbf{e}^j\mathbf{B}}{2} &= \frac{\mathrm{i}}{2}[B^j\mathbf{E} + (\mathbf{B}\wedge\mathbf{e}^j)\mathbf{E} - \mathbf{E}B^j - \mathbf{E}(\mathbf{e}^j\wedge\mathbf{B})] \\ &= \boldsymbol{i}(\mathbf{E}\wedge\mathbf{B}\wedge\mathbf{e}^j) = (\boldsymbol{i}\mathbf{E}\wedge\mathbf{B})\cdot\mathbf{e}^j = -(\mathbf{E}\times\mathbf{B})\cdot\mathbf{e}^j = -\mathbf{S}\cdot\mathbf{e}^j,\end{aligned} \tag{210}$$

which gives the energy flux along the direction of $\mathbf{e}^j$. The last term can be calculated using $\mathbf{v}\mathbf{e}^j = 2v^j - \mathbf{e}^j\mathbf{v}$, and we have

$$\begin{aligned}\frac{\mathbf{E}\mathbf{e}^j\mathbf{E} + \mathbf{B}\mathbf{e}^i\mathbf{B}}{2} &= E^j\mathbf{E} + B^j\mathbf{B} - \frac{\mathbf{e}^j\mathbf{E}^2 + \mathbf{e}^j\mathbf{B}^2}{2} \\ &= \left(E^jE^k + B^jB^k - \frac{\mathbf{E}^2 + \mathbf{B}^2}{2}\delta^{jk}\right)\mathbf{e}_k = \tau^{jk}\mathbf{e}_k.\end{aligned} \tag{211}$$

The quantity τ^{jk} is the Maxwell stress tensor.

Finally, using all the above expressions in equation (204), we obtain

$$-\partial_t u - \partial_t\mathbf{g} - \partial_j S^j + \partial_j\tau^{jk}\mathbf{e}_k = \mathbf{E}\cdot\mathbf{j}_0 + \rho_0\mathbf{E} + \mathbf{j}_0\times\mathbf{B}, \tag{212}$$

The scalar part of equation (212) is the Poynting theorem and the vector part is the momentum version of the Poynting theorem [39]. If we define the paravectors $\mathcal{U}$ and $\mathcal{U}^j$

($j = 1, 2, 3$),

$$\mathcal{U} = u + \mathbf{g}, \tag{213}$$

$$\mathcal{U}^j = S^j - \tau^{jk}\mathbf{e}_k, \tag{214}$$

we can write equation (212) as

$$\partial_t \mathcal{U} + \partial_j \mathcal{U}^j + f = 0. \tag{215}$$

References

[1] Tomonaga S 1997 *The Story of Spin trans T Oka* (Chicago, IL: University of Chicago Press)
[2] Pauli W 1925 *Z. Phys.* **31** 373–85
[3] Uhlenbeck G E and Goudsmit S A 1925 *Naturwissenschaften* **13** 953–4
[4] Heisenberg W and Jordan P 1926 *Z. Phys.* **37** 263–77
[5] Pauli W 1927 *Z. Phys.* **43** 601–23
[6] Dirac P A M 1928 *Proc. R. Soc.* A **117** 610–24
[7] Schwabl F 2005 *Advanced Quantum Mechanics* 3rd edn (Berlin: Springer)
[8] Thaller B 1992 *The Dirac Equation* (Berlin: Springer)
[9] Dyson F 1972 *Bull. Am. Math. Soc.* **78** 635–52
[10] Algebra of physical space *Wikipedia* http://en.wikipedia.org/wiki/Algebra_of_physical_space (accessed: 2 April 2016)
[11] Lounesto P 2001 *Clifford Algebras and Spinors* 2nd edn (Cambridge: Cambridge University Press)
[12] Porteous I 1995 *Clifford Algebras and the Classical Groups* (Cambridge: Cambridge University Press)
[13] Vaz J Jr and Rocha R Jr 2016 *An Introduction to Clifford Algebras and Spinors* (Oxford: Oxford University Press)
[14] Goldstein H, Poole C P Jr and Safko J L 2001 *Classical Mechanics* 3rd edn (Harlow: Pearson)
[15] Hestenes D 1973 *J. Math. Phys.* **14** 893–905
[16] Hestenes D 1975 *J. Math. Phys.* **16** 556–72
[17] Baylis W E 1999 *Phys. Rev.* A **60** 785–95
[18] Baylis W E 1996 The paravector model of spacetime *Clifford (Geometric) Algebras with Applications in Physics, Mathematics and Engineering* (Boston, MA: Birkhauser) pp 237–52
[19] Baylis W E 2004 Applications of Clifford algebras in physics *Lectures on Clifford (Geometric) Algebras and Applications* ed R Ablamowicz and G Sobczyk (Boston, MA: Birkhauser) pp 91–133
[20] Baylis W E 2004 Geometry of paravector space with applications to relativistic physics *Computational Noncommutative Algebra with Applications* ed J Byrnes (Dordrecht: Kluwer) pp 363–87
[21] Baylis W E 2004 Quantum/classical interface: a geometric approach from the classical side *Computational Noncommutative Algebra with Applications* ed J Byrnes (Dordrecht: Kluwer) pp 127–54
[22] Baylis W E and Keselica J D 2012 *Adv. Appl. Clifford Algebras* **22** 537–61
[23] Vaz J Jr 2013 *Int. J. Theor. Phys.* **52** 1440–54
[24] Ritchie B and Weatherford C A 2012 *Int. J. Quantum Chem.* **112** 3722–8
[25] Daviau C 2012 *Adv. Appl. Clifford Algebras* **22** 611–23
[26] Greiner W 2000 *Relativistic Quantum Mechanics: Wave Equations* 3rd edn (Berlin: Springer)
[27] Doran C and Lasenby A 2003 *Geometric Algebra for Physicists* (Cambridge: Cambridge University Press)
[28] Takahashi Y 1982 *Phys. Rev.* D **26** 2169–71
[29] Boudet R 1985 *J. Math. Phys.* **26** 718–24
[30] Ahluwalia D V and Grumiller D 2005 *J. Cosmol. Astropart. Phys.* JCAP07(2005)012
[31] da Rocha R and Rodrigues W A Jr 2006 *Mod. Phys. Lett.* A **21** 65
[32] da Rocha R and and Hoff da Silva J M 2007 *J. Math. Phys.* **48** 123517

[33] de Oliveira E C, Rodrigues W A Jr and Vaz J Jr 2014 *Int. J. Theor. Phys.* **53** 4381
[34] da Rocha R, Bernardini A E and Hoff da Silva J M 2011 *J. High Energy Phys.* JHEP04(2011)110
[35] Ahluwalia D V and Nayak A C 2014 *Int. J. Mod. Phys.* D **23** 1430026
[36] Wigner E P 1964 Unitary representations of the inhomogeneous Lorentz group including reflections *Group Theoretical Concepts in Elementary Particle Physics* ed F Gursey (New York: Gordon and Breach) pp 37–80
[37] Houtappel R M F, van Dam H and Wigner E P 1965 *Rev. Mod. Phys.* **37** 595
[38] Felix F R 2016 Aplicações da álgebra do espaço físico em mecânica quântica *MSc Dissertation* IMECC-UNICAMP (in Portuguese, unpublished)
[39] Panofsky W K H and Phillips M 1962 *Classical Electricity and Magnetism* 2nd edn (Reading, MA: Addison-Wesley)