A quaternion representation of the Lorentz group for classical physical applications

1991 J. Phys. A: Math. Gen. 24 3245

(http://iopscience.iop.org/0305-4470/24/14/013)

J. Phys. A: Math. Gen. 24 (1991) 3245–3254. Printed in the UK

A quaternion representation of the Lorentz group for classical physical applications

I Abonyi†, J F Bito‡ and J K Tar‡

† Department of Theoretical Physics, Roland Eötvös University, H-1088 Budapest, Puskin u. 5-7, Hungary
‡ Tungsram Trading House Co Ltd, Centre for Robotics, H-1340 Budapest, Váci út 77, Hungary

Received 3 January 1991, in final form 3 April 1991

Abstract. A special quaternion representation is constructed for a pair of relativistic vectors and skew-symmetric tensors on the basis of the group theory of Lorentz transformations. The construction has considerable advantages over the conventional vector–tensor description. It is pointed out that pairs of Minkowski vectors as well as certain scalars and skew-symmetric tensors can also be interpreted as simple components of more complex physical quantities, each of them expressed by a single quaternion. As an example a concise relativistic quaternion formulation of Maxwell's electrodynamics is presented. The relativistic covariance can be maintained even for the existence of magnetic monopoles.

1. Introduction

Though in his 'Treatise' J C Maxwell applied three-dimensional (3D) vector representation to formulate electrodynamics, he also mentioned a possible 3D quaternion application to describe vectors [1]. This application utilized the fact that both scalar and vector products of 3D vectors automatically appear in the quaternion products. Since neither the concept nor the significance of the relativistic covariance was realized at that, Maxwell used the VQ symbol to refer to the 'vector part' of the quaternion Q, while SQ referred to the 'scalar part' of Q. Since this notation does not provide plausible advantages over the 3D vector formulation from the classical point of view, quaternions faded into oblivion for physicists. While their renaissance began with quantum theory in the form of Pauli's spinormatrices or related to more general considerations in quantum physics [2, 3], technical applications seem only to have spread in our time [4].

In robotics, the convenient behaviour of quaternions for $O(3)^+$ transformations is utilized to describe kinematic structure and to compute rotations of robots' arms [5]. Analysis of the cause of the main computational benefits lead us to investigate the possibilities of the quaternion representation of the Lorentz group. Based on the fact that the $O(3)^+$ group is a subgroup of the relativistic Lorentz group, this representation is successfully generalized to the whole continuous Lorentz group. In contrast to the more general approach published in [2] and [3], we considered a power series formulation of the most general transformations generated by a 'mixture' of pure $O(3)^+$ and pure Lorentzian generators.

The main benefits of the proposed method can be found in a closed analytic formulation of arbitrary Lorentz transformations of arbitrary pairs of Minkowski vectors, general skew-symmetric and special symmetric tensors. Instead, quaternionic power series formulae contain the simple complex functions $\sinh z$ and $\cosh z$, which can be easily computed.

From the educational point of view it is quite interesting that strict inner correspondence can be found between relativistic electrodynamics and the quaternion representation. Each traditionally used equation can be obtained as a result of plausible algebraic manipulations with quaternions.

2. Quaternion representation for the Lorentz group

Using special units, i.e. $c=1$ for the velocity of light *in vacuo*, the covariant representation of the metric tensor of the Minkowski space can be written as

$$\begin{aligned} g_{ij} &= 1 \qquad \text{for } i=j=1,2,3 \\ g_{44} &= -1 \\ g_{ij} &= 0 \qquad \text{otherwise.} \end{aligned} \tag{2.1}$$

The Lorentz matrices L of the $x^{i} = L^{i}_{j} x^{j}$ transformation for contravariant 4D relativistic vectors can be expressed as exponential series in which the exponents are linear combinations of the generators $\{U^i,\ V^i \mid i=1,2,3\}$ of this Lie group. The skew-symmetric 4×4 matrices U^i generate the whole subgroup $O(3)^+$, while the symmetric matrices V^i generate pure Lorentz transformations. The generators obey the commutation rules

$$[U^i, U^j] = U^u \varepsilon_{uij} \qquad [V^i, V^j] = -U^u \varepsilon_{uij} \qquad [U^i, V^j] = V^u \varepsilon_{uij} \tag{2.2}$$

where ε_{uij} is the Levi-Cività symbol. The quanternion representation of the group will be constructed via the representation of this Lie algebra. In this representation the quaternions serve as a complex 4D linear vector space. Introducing the multiplication rules by the relations

$$\begin{aligned} &K^0K^i = K^iK^0 = K^i \qquad K^0K^0 = K^0 \equiv I \\ &K^iK^j = -I\delta_{ij} + K^u\varepsilon_{uij} \qquad (i=1,2,3) \\ &(\alpha Q)(\beta P) := (\alpha\beta)(QP) \end{aligned} \tag{2.3}$$

where α, β are arbitrary complex numbers, quaternions also form an algebraic structure. In order to find an appropriate representation for (2.2) we decompose this complex 4D linear vector space into a real 8D linear space on the following basis:

$$\{I, K^i, H^0 \equiv J := \mathrm{i}I, H^j := \mathrm{i}K^j\} \tag{2.4}$$

where i denotes the imaginary unit. It is easy to see that there is more than one possibility to represent (2.2) by the use of quaternions. For instance, the representations

$$U^i \to D(U^i) := \tfrac{1}{2}K^i \qquad V^i \to D(V^i) := \tfrac{1}{2}H^i \tag{2.5}$$

$$U^i \to D(U^i) := \tfrac{1}{2}[K^i, \cdot\,] \qquad V^i \to D(V^i) := \tfrac{1}{2}[H^i, \cdot\,] \tag{2.6}$$

$$U^i \to D(U^i) := \tfrac{1}{2}[K^i, \cdot\,] \qquad V^i \to D(V^i) := \tfrac{1}{2}\{H^i, \cdot\,\} \tag{2.7}$$

equally satisfy the laws of the Lie algebra (2.2). (Here the symbol $[\cdot, \cdot]$ denotes commutators while $\{\cdot, \cdot\}$ corresponds to anticommutators.) Using one of the equations

(2.5)-(2.7), with arbitrary real numbers a_s and b_s, the quaternion representation of the Lorentz matrices L can be defined as

$$D[\exp(a_sU^s+b_sV^s)]:=\exp[a_sD(U^s)+b_sD(V^s)]. \tag{2.8}$$

Due to the same structure of the exponentials this representation has the feature $D(QP)=D(Q)D(P)$, i.e. it serves as a correct representation of the Lorentz group. According to the rules in (2.3), it is easy to express the exponentials pertaining to (2.5). Introducing the complexes $\{\alpha_j:=a_j+\mathrm{i}b_j\,|\,j=1,2,3\}$ and the generators $G:=\alpha_iK^i$, we get

$$\begin{aligned}&G^0:=I \qquad G^2=\alpha_i\alpha_jK^iK^j=\alpha_i\alpha_j(-I\delta_{ij}+K^u\varepsilon_{uij})=-I\alpha_i\alpha_i:=\beta I\\ &G^3=(\sqrt{\beta})^3G/(\sqrt{\beta}) \qquad G^4=(\sqrt{\beta})^4I \qquad \ldots .\end{aligned} \tag{2.9}$$

Hence for an arbitrary real group parameter 's' we have

$$\exp(\tfrac{1}{2}s\alpha_iK^i)=\cosh(\tfrac{1}{2}s\sqrt{\beta})I+\sinh(\tfrac{1}{2}s\sqrt{\beta})(\alpha_i/\sqrt{\beta})K^i. \tag{2.10}$$

We note that a unique branching cut on the complex plane must be defined for $\sqrt{\beta}$ in order to get correct results. For the special case $b_i\equiv 0$ when $\sqrt{\beta}$ is purely imaginary we obtain the usual quaternion representation of the $\mathrm{O}(3)^+$ group. For $\sqrt{\beta}=\mathrm{i}$ it corresponds to 's', given in radians, describing an $\mathrm{O}(3)^+$ rotation around the unit vector axis $e_i:=\mathrm{i}a_i/\sqrt{\beta}$, and

$$Q(s,\boldsymbol{e})=\cos(\tfrac{1}{2}s)I+e_i\sin(\tfrac{1}{2}s)K^i. \tag{2.11}$$

For the special case of pure Lorentz transformations $a_i\equiv 0$, (2.10) gives

$$Q(s,\boldsymbol{e})=\cosh(\tfrac{1}{2}s)I+\mathrm{i}\sinh(\tfrac{1}{2}s)e_uK^u. \tag{2.12}$$

Here $e_i:=b_i/\sqrt{\beta}$ corresponds to the direction of the relative velocity of the frames connected by the transformation L, while s serves as the velocity parameter. Now we consider the representation given in (2.7). For an arbitrary quaternion Q the most general generator $\boldsymbol{G}$ has the effect

$$\boldsymbol{G}Q=\tfrac{1}{2}a_i[K^i,Q]+\tfrac{1}{2}\mathrm{i}b_i\{K^i,Q\}=\tfrac{1}{2}(a_i+\mathrm{i}b_i)K^iQ-\tfrac{1}{2}Q(a_i-\mathrm{i}b_i)K^i:=\tfrac{1}{2}(zQ-Qz^*).$$

Considering the powers $\boldsymbol{G}^0$, $\boldsymbol{G}^1$, $\boldsymbol{G}^2$, $\boldsymbol{G}^3$ we get

$$\begin{aligned}&\boldsymbol{G}^0Q:=IQ \qquad \boldsymbol{G}^1Q=\tfrac{1}{2}(zQ-Qz^*)\\ &\boldsymbol{G}^2Q=(\tfrac{1}{2})^2(z^2Q-2zQz^*+Qz^{*2})\\ &\boldsymbol{G}^3Q=(\tfrac{1}{2})^3(z^3Q-3z^2Qz^*+3zQz^{*2}-Qz^{*3})\end{aligned} \tag{2.13}$$

where the members of the binomial series can be recognized in the role of the coefficients. Therefore the nth term in the power series of $\exp(\boldsymbol{G}s)$ has the form

$$\boldsymbol{G}^nQ/(n!)=\sum_{m=0}^{n} z^mQ(-z^*)^{n-m}/[m!(n-m)!]. \tag{2.14}$$

Summing up (2.14) for $n:0\to\infty$ and rearranging the terms in the series we can find

$$\exp(s\boldsymbol{G})Q=\exp(\tfrac{1}{2}sz)Q\exp(-\tfrac{1}{2}sz^*). \tag{2.15}$$

3. Quaternion representation of vectors and tensors

Considering the infinitesimal transformation $\mathrm{d}L$ for an arbitrary contravariant 4D vector we get

$$\mathrm{d}L\,x=\mathrm{d}s(a_m U^m+b_m V^m)[\boldsymbol{r},t]=\mathrm{d}s[\boldsymbol{a}\times\boldsymbol{r}+\boldsymbol{b}t,\boldsymbol{br}]. \tag{3.1}$$

Let $x=[\boldsymbol{r},x_0]$, $y=[\boldsymbol{y},y_0]$ be two contravariant 4D vectors. These two independent vectors can easily be represented by the quaternion

$$q_{\mathrm{con}}:=(x_0-\mathrm{i}y_0)K^0+(\mathrm{i}r_s+y_s)K^s. \tag{3.2}$$

(The 'con' subscript refers to the word 'contravariant'.) Calcultating the quaternion representation $\mathrm{d}\boldsymbol{q}:=\mathrm{d}s\,\boldsymbol{G}\boldsymbol{q}$ for (3.1) we get

$$\begin{aligned} &\mathrm{d}x_0=\mathrm{d}s\,\boldsymbol{br} && \mathrm{d}y_0=\mathrm{d}s\,\boldsymbol{by}\\ &\mathrm{d}\boldsymbol{x}=\mathrm{d}s(\boldsymbol{a}\times\boldsymbol{r}+x_0\boldsymbol{b}) && \mathrm{d}\boldsymbol{v}=\mathrm{d}s(\boldsymbol{a}\times\boldsymbol{y}+y_0\boldsymbol{b}). \end{aligned} \tag{3.3}$$

This means, that (3.2) is a quaternion representation of two independent contravariant 4D Minkowski vectors with the tranformation law of

$$q'_{\mathrm{con}}=\exp(\tfrac{1}{2}sz)q_{\mathrm{con}}\exp(-\tfrac{1}{2}sz^*). \tag{3.4}$$

Considering now the transformation law for the covariant components $x_i:=g_{ij}x^j$, according to (2.1), the appropriate quaternion representation is

$$q_{\mathrm{cov}}:=(-x_0+\mathrm{i}y_0)K^0+(\mathrm{i}r_s+y_s)K^s. \tag{3.5}$$

The appropriate transformation law can be obtained from (3.3) taking $\boldsymbol{b}\to-\boldsymbol{b}$, i.e.

$$q'_{\mathrm{cov}}=\exp(\tfrac{1}{2}sz^*)q_{\mathrm{cov}}\exp(-\tfrac{1}{2}sz). \tag{3.6}$$

Let R_{cov} and Q_{con} be representations for two covariant and two contravariant Minkowski vectors respectively. The transformation laws for the products $R_{\mathrm{cov}}Q_{\mathrm{con}}$ and $Q_{\mathrm{con}}R_{\mathrm{cov}}$ turn out to be

$$\begin{aligned} R'_{\mathrm{cov}}Q'_{\mathrm{con}}&=\exp(\tfrac{1}{2}sz^*)R_{\mathrm{cov}}\exp(-\tfrac{1}{2}sz)\exp(\tfrac{1}{2}sz)Q_{\mathrm{con}}\exp(-\tfrac{1}{2}sz^*)\\ &=\exp(\tfrac{1}{2}sz^*)R_{\mathrm{cov}}Q_{\mathrm{con}}\exp(-\tfrac{1}{2}sz^*) \end{aligned}$$

and

$$\begin{aligned} Q'_{\mathrm{con}}R'_{\mathrm{cov}}&=\exp(\tfrac{1}{2}sz)Q_{\mathrm{con}}\exp(-\tfrac{1}{2}sz^*)\exp(\tfrac{1}{2}sz^*)R_{\mathrm{cov}}\exp(-\tfrac{1}{2}sz)\\ &=\exp(\tfrac{1}{2}sz)Q_{\mathrm{con}}R_{\mathrm{cov}}\exp(-\tfrac{1}{2}sz). \end{aligned} \tag{3.7}$$

In order to calculate the components of (3.7) we rewrite (3.2) as

$$(x_0-\mathrm{i}y_0)K^0+(\mathrm{i}r_s+y_s)K^s=x_0I+\mathrm{i}r_sK^s-\mathrm{i}(y_0+\mathrm{i}y_sK^s)=x-\mathrm{i}y. \tag{3.8}$$

Therefore, (3.7) can be written as a composition of rather simple terms:

$$(A_{\mathrm{con}}-\mathrm{i}B_{\mathrm{con}})(C_{\mathrm{cov}}-\mathrm{i}D_{\mathrm{cov}})=\mathrm{A}_{\mathrm{con}}C_{\mathrm{cov}}-B_{\mathrm{con}}D_{\mathrm{cov}}-\mathrm{i}(A_{\mathrm{con}}D_{\mathrm{cov}}+B_{\mathrm{con}}C_{\mathrm{cov}}). \tag{3.9}$$

For the simple $A_{\mathrm{con}}C_{\mathrm{cov}}$ term we get

$$\begin{aligned} &(A_0I+\mathrm{i}A_sK^s)(-C_0I+\mathrm{i}C_uK^u)\\ &\qquad=(-A_0C_0+A_sC_s)I+\mathrm{i}(A_0C_s-C_0A_s)K^s-K^s\varepsilon_{suw}A_uC_w. \end{aligned} \tag{3.10}$$

The scalar part of (3.10) corresponds to the scalar product (A, C) while the other parts are equal to the six independent components of the antisymmetrized direct product $A\cdot C-C\cdot A$ tensor. Therefore each component in (3.9) can be easily interpreted. Furthermore, the given transformation laws guarantee that appropriate relativistically covariant quaternion products can be constructed: a (scalar+tensor)*(scalar+tensor) product transforms as (scalar+tensor) quaternion, while the vector*(scalar+tensor) product transforms as a vector quaternion. It is also clear from (3.10) that there is a general possibility to represent antisymmetric tensors as quaternions. We note that no general way exists to to represent arbitrary symmetric tensors with quaternions.

4. Some remarks on the possible ease of computation

In this section we intend to show a well-defined special case in which the use of quaternions can afford considerable ease: computation of the result of numerous consecutive $O(3)^+$ rotations. Such a situation is typical in robotics where the different arm sections rotating with respect to each other form a kinetic chain. This case may also have some relativistic interest. Since the rotating sections correspond to non-inertial frames, their 'absolute rotation' can be detected and used for control by optical methods (so-called Sagnac encoders).

For the computation of the result of two consecutive Lorentz transformations the product of two 4×4 matrices is to be calculated. For each of the 16 matrix elements four real multiplications and three real additions occur which means 64 real multiplications and 48 real additions.

In general, the situation is not at all better for quaternions. According to the rule in (2.3), the product of two complex quaternions consists of the components

$$(A_0I+A_sK^s)(B_0I+B_tK^t)=(A_0B_0-A_sB_s)I+(A_0B_s+B_0A_s+\varepsilon_{sut}A_uB_t)K^s \tag{4.1}$$

which means four complex product and three complex additions in the scalar component and $1+1+2=4$ complex multiplications with three complex additions for each of the three 'space' components. The sum is 16 complex multiplciations and 12 complex additions. Taking into account the real and imaginary parts, each complex multiplication means four real multiplications and two real additions, therefore in the general case 64 real multiplications and 56 real additions are necessary when quaternions are used.

However, the situation is very much better in the case of the $O(3)^+$ subgroup. According to (2.11) only real components occur now, and in this case the 16 multiplications with the 12 additions afford considerable ease over the 64 multiplications and 48 additions.

5. Connection with spinor representations

As is well known, a strict connection between spinors and the elements of the Lorentz group can be established via elementary considerations as follows. For Lorentz transformations the vectors of the light cone having $x_\mu x^\mu=0$ are transformed to each other. Because each component of such vectors can be so scaled that the result has a space component of half length, the most general Lorentz transformation realizes a transformation of a 2D sphere into itself. Furthermore, as a 2D surface, this ball can be mapped onto the complex plane in different ways, e.g.

$$\begin{aligned} z_1 &= (x_1 - \mathrm{i}x_2)/(\tfrac{1}{2} - x_3) \\ z_2 &= (x_3 - \mathrm{i}x_1)/(\tfrac{1}{2} - x_2) \\ z_3 &= (x_2 - \mathrm{i}x_3)/(\tfrac{1}{2} - x_1) \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} x_3 &= \tfrac{1}{2}(z_1 z_1^* - 1)/(z_1 z_1^* + 1) \\ x_1 &= \mathrm{Re}\, z_1/(z_1 z_1^* + 1) \\ x_2 &= -\mathrm{Im}\, z_1/(z_1 z_1^* + 1) \qquad \text{etc} \end{aligned} \tag{5.2}$$

and

$$z_1 = (z_3 - \mathrm{i})/(z_3 + \mathrm{i}) \qquad z_3 = (z_2 - \mathrm{i})/(z_2 + \mathrm{i}) \tag{5.3}$$

where the half length of the space component has been utilized. Regarding $O(3)^+$, each rotation can be obtained as the result of consecutive rotations around the axes x_3, x_2 and x_1 with different angles $\{q_i \mid i = 1, 2, 3\}$. In addition, each pure Lorentz boost can easily be expressed by rotating a special boost around a given axis. For instance, for rotations and boosts around/according to the x_3 direction z_1 has a very simple transformation as

$$\begin{aligned} z_1' &= \exp(-\mathrm{i}q_3)z_1 \qquad \text{for rotations} \\ z_1' &= \exp(q)z_1 \qquad \text{for boosts} \end{aligned} \tag{5.4}$$

etc. According to (5.3) each element of the Lorentz group can be expressed with the transformation law for z_1 having a general fractional form of

$$z_1' = (\alpha z_1 + \beta)/(\mu z_1 + \delta) \tag{5.5}$$

where the α, β, μ, δ complexes are not determined unambiguously. Observing that resulting coefficients of consecutive fractional transformations can be expressed as matrix products with the matrices $A_{11} = \alpha$, $A_{12} = \beta$, $A_{21} = \mu$ and $A_{22} = \delta$ it is expedient to impose the det $A = 1$ restriction which makes A unambiguous apart from a ± 1 factor. The unimodular 2×2 complex matrices form the $\mathrm{Sl}(2, C)$ group having the same Lie algebra as the Lorentz group. The 2D complex linear vector space on which the $\mathrm{Sl}(2, C)$ matrices operate is called the space of spinors. According to (5.4) it is easy to see that the unitary subgroup of $\mathrm{Sl}(2, C)$, called SU(2), just corresponds to the $O(3)^+$ subgroup of the Lorentz group.

Regarding the restriction of det $A = 1$, by the use of (5.4) we can observe that spinor representations must have a $\frac{1}{2}$ factor in the exponents of the appropriate matrix elements. For the $O(3)^+$ subgroup it has the consequence that rotations with an angle of 2π are described by -1 and due to the requirement of continuity the $L \to \pm A$ correspondence cannot be made unambiguous by dropping 'one-half' of the $\mathrm{Sl}(2, C)$ group. Therefore, rigorously speaking, $\mathrm{Sl}(2, C)$ with spinors do not represent the Lorentz group (it represents only its Lie algebra), and thus it should be excluded when constructing representations of the Lorentz group.

However, it is the revelation of quantum physics that the essential symmetry is not the Lorentz group but the broader $\mathrm{Sl}(2, C)$ itself. Due to exhaustive theoretical investigations it has also become clear that the scalar and spinor representations have a fundamental significance: each finite-dimensional representation of $\mathrm{Sl}(2, C)$ can be constructed by the use of scalars and direct products of spinor fields. Therefore, tensorial representations can also be constructed with spinors and spinors find extensive applications in physics, e.g. see [6–8].

Regarding the connection with quaternions, spinor representations seem to be similar to representations given in (2.5) and (2.10), because in both cases 2π rotations are described by the -1 transformation. Furthermore, due to the rule for matrix exponentials that

$$\det[\exp(tM)] = \exp(t \operatorname{Tr} M) \tag{5.6}$$

to satisfy the det $A = 1$ restriction, traceless Hermitian generators have to be found to represent $\mathrm{Sl}(2, C)$. The appropriate linearly independent matrices are Pauli's spin matrices $\{\sigma_i | i = 1, 2, 3\}$ with such a multiplication rule that $\{\sigma_0, -\mathrm{i}\sigma_j\}$ exactly serves as a matrix representation of the multiplication rule for quaternions given in (2.3) (σ_0 is the unit matrix). Another representation can be obtained for Minkowski vectors as Hermitian 2×2 complex matrices,

$$x := x_\mu \sigma^\mu \qquad x' = AxA^+ \tag{5.7}$$

where $x_\mu x^\mu = \det(x)$. This representation is rather similar to the transformation rule given in (3.4).

6. Maxwell's electrodynamics with quaternions

Here Maxwell's electrodynamics will be discussed with quaternions. In order to avoid the use of the μ_0, ε_0 constants for vacuum, we use a CGS-like system of units, with the $c = 1$ modification. First we introduce the Grad 'gradient' and Γ 'potential' quaternions as

$$\mathrm{Grad}_{\mathrm{cov}} := \partial_t I + \mathrm{i}K^m \partial_m \qquad \Gamma_{\mathrm{con}} := \Phi I + \mathrm{i}K^m A_m. \tag{6.1}$$

The ∂_t symbol stands for the partial time derivative, while ∂_m ($m = 1, 2, 3$) correspond to 'space derivatives'. The $\mathrm{Grad}_{\mathrm{cov}}\Gamma_{\mathrm{con}}$ product has the following components:

$$\begin{aligned} \mathrm{Grad}_{\mathrm{cov}}\Gamma_{\mathrm{con}} &= (\partial_t \Phi + \operatorname{div} A) I - (\operatorname{rot} A)_m K^m + (\partial_t A + \operatorname{grad} \Phi)_m \,\mathrm{i}K^m \\ &\equiv \mathrm{Gauge}\, I - B_m K^m - E_m \,\mathrm{i}K^m = \mathrm{Gauge}\, I + F_{\mathrm{cov,con}}. \end{aligned} \tag{6.2}$$

Equation (6.2) consists of the invariant scalar 'Gauge' and the quaternion representation of the skew-symmetric field tensor. In similar way we can see that the

$$\mathrm{Grad}_{\mathrm{con}}\mathrm{Grad}_{\mathrm{cov}} = (-\partial_t^2 + \partial_m^2) \equiv \Box \tag{6.3}$$

product yields the pure relativistic scalar d'Alembertian operator $\Box$. Let us now make use of the associativity of the $\mathrm{Grad}_{\mathrm{con}}\mathrm{Grad}_{\mathrm{cov}}\Gamma_{\mathrm{con}}$ product, which transforms as a vector. We get

$$\Box\Gamma_{\mathrm{con}} = \mathrm{Grad}_{\mathrm{con}}\mathrm{Gauge} + \mathrm{Grad}_{\mathrm{con}} F_{\mathrm{cov,con}}. \tag{6.4}$$

According to Maxwell's usual equations we obtain

$$\mathrm{Grad}_{\mathrm{con}} F_{\mathrm{cov,con}} = -4\pi J_{\mathrm{con}} \tag{6.5}$$

where $J_{\mathrm{con}} = \sigma I + \mathrm{i}j_m K^m$ is the electric charge and the current density vector. Putting (6.5) into (6.4) we obtain the usual wave equations for the potential Γ. The conservation law of electric charge is derived by multiplying both sides of (6.5) by $\mathrm{Grad}_{\mathrm{cov}}$, and observing that on the left-hand side we have the term $\Box F_{\mathrm{cov,con}}$, the scalar part of which is zero. All these lead to the continuity equation of the electric charge:

$$\mathrm{Grad}_{\mathrm{cov}} J_{\mathrm{con}} = 0. \tag{6.6}$$

To obtain the conservation law for the energy-momentum let us consider the real scalar and the imaginary 'space' complex parts of the $-4\pi F_{\mathrm{cov,con}} J_{\mathrm{cov}}$ product! It is

$$4\pi(\boldsymbol{Ej})I - 4\pi\mathrm{i}(\boldsymbol{E}\sigma - \boldsymbol{B}\times\boldsymbol{j})_m K^m \tag{6.7}$$

where it is easy to recognize the power and force density vector of the field-charge interaction, respectively:

$$\begin{aligned} 4\pi(\boldsymbol{Ej})I - 4\pi\mathrm{I}(\boldsymbol{E}\sigma - \boldsymbol{B}\times\boldsymbol{j})_m K^m &= F_{\mathrm{cov,con}}\mathrm{Grad}_{\mathrm{cov}} F_{\mathrm{con,cov}} \\ &= -[\partial_t \tfrac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) + \mathrm{div}(\boldsymbol{E}\times\boldsymbol{B})]I \\ &\quad + [\partial_t(\boldsymbol{E}\times\boldsymbol{B}) + \boldsymbol{B}\times\mathrm{rot}\,\boldsymbol{B} - \boldsymbol{B}\,\mathrm{div}\,\boldsymbol{B} + \boldsymbol{E}\times\mathrm{rot}\,\boldsymbol{E} - \boldsymbol{E}\,\mathrm{div}\,\boldsymbol{E}]_m \mathrm{i}K^m. \end{aligned} \tag{6.8}$$

Taking into account that $\varepsilon_{ijk}\varepsilon_{ils} = \delta_{jl}\delta_{ks} - \delta_{js}\delta_{kl}$, we have

$$\boldsymbol{B}\times\mathrm{rot}\,\boldsymbol{B} - \boldsymbol{B}\,\mathrm{div}\,\boldsymbol{B} + \boldsymbol{E}\times\mathrm{rot}\,\boldsymbol{E} - \boldsymbol{E}\,\mathrm{div}\,\boldsymbol{E} = \mathrm{div}[\tfrac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) - \boldsymbol{B}\cdot\boldsymbol{B} - \boldsymbol{E}\cdot\boldsymbol{E}] := \mathrm{div}\,4\pi\mathsf{T}.$$

where T is the Maxwellian momentum density tensor. Therefore the components of (6.8) contain the energy density $u := \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2)/4\pi$, the energy current density vector $\boldsymbol{S} := \boldsymbol{E}\times\boldsymbol{B}/4\pi$ (which is equal to the momentum density vector), and the Maxwellian momentum density tensor T.

Now the whole physical interpretation is as follows. The equation of the real parts give

$$4\pi(\boldsymbol{Ej}) = -[\partial_t \tfrac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) + \mathrm{div}(\boldsymbol{E}\times\boldsymbol{B})] \tag{6.9}$$

which evidently is the equation of continuity of the field energy density $(1/8\pi)(\boldsymbol{E}^2 + \boldsymbol{B}^2)$, the sources and sinks are represented by the Poynting vector $(1/4\pi)$ $(\boldsymbol{E}\times\boldsymbol{B})$ and the Joule heat production by $\boldsymbol{Ej}$. The equation between the imaginary parts leads to

$$\boldsymbol{E}\sigma + \boldsymbol{j}\times\boldsymbol{B} = -(1/4\pi)\{\partial_t(\boldsymbol{E}\times\boldsymbol{B}) + \mathrm{div}\,\mathsf{T}\} \tag{6.10}$$

telling us that the force acting on the charge density σ and current density $\boldsymbol{j}$ comes from the change in time of the momentum density of the electromagnetic field and the divergence of the Maxwellian stresses.

As we noted in section 3, it is not possible to represent the 10 independent components of a general symmetric tensor by the use of eight independent quaternion coefficients. The relativistic energy and momentum density tensor $\mathsf{T}_{\alpha\beta}$ is symmetric, but it can be completely constructed out of the six independent components of $\boldsymbol{E}$ and $\boldsymbol{B}$. Thereofre the relativistically covariant expression $\partial_\alpha T^{\alpha\beta}$ has a quaternion representation. The relativistically covariant quaternion product

$$F_{\mathrm{cov,con}} F_{\mathrm{cov,con}} = (\boldsymbol{E}^2 - \boldsymbol{B}^2) + \mathrm{i}2(\boldsymbol{EB}) \tag{6.11}$$

corresponds to the two relativistic scalar invariants of the electromagnetic field (e.g. [9]). Though the special product

$$\tfrac{1}{2}F_{\mathrm{cov,con}} F_{\mathrm{con,cov}} = -uI + \mathrm{i}(\boldsymbol{E}\times\boldsymbol{B})_s K^s = -\tfrac{1}{2}F_{\mathrm{con,cov}} F_{\mathrm{cov,con}} \tag{6.12}$$

does not tranform as a vector (the inner $\exp(-\frac{1}{2}sz^*)\exp(\frac{1}{2}sz)$ products do not provide the unit quaternion I in a general case in the appropriate equations similar to (3.7)), it is also worthy of note. It represents the last column (and last row) of $T_{\alpha\beta}$.

7. Extension to magnetic monopoles

Now we consider the case in which div $\boldsymbol{B} \neq 0$! In the most general case the Γ-potential can be written as

$$\Gamma_{\text{cov}} = (-\Phi + \mathrm{i}\phi)I + \mathrm{i}A_m K^m + a_m K^m = \Gamma^{\text{E}}_{\text{cov}} - \mathrm{i}\Gamma^{\text{M}}_{\text{cov}}. \tag{7.1}$$

The relativistically covariant $\text{Grad}_{\text{con}}\Gamma_{\text{cov}}$ product yields the relations

$$\begin{aligned}
&\text{gauge} = \partial_t \Phi + \text{div}\, \boldsymbol{A} - \mathrm{i}(\partial_t \phi + \text{div}\, \boldsymbol{a}) \\
&\boldsymbol{E} = -\text{grad}\, \Phi - \partial_t \boldsymbol{A} + \text{rot}\, \boldsymbol{a} \\
&\boldsymbol{B} = \text{rot}\, \boldsymbol{A} + \partial_t \boldsymbol{a} + \text{grad}\, \phi \\
&-B_m K^m - E_m\, \mathrm{i}K^m = F_{\text{cov,con}}.
\end{aligned} \tag{7.2}$$

That is to say the electric charge density $\sigma^{\text{E}} := \text{div}\, \boldsymbol{E}/4\pi$ remains unchanged, while a non-zero 'magnetic charge density',

$$\sigma^{\text{M}} := \text{div}\, \boldsymbol{B}/4\pi = (\partial_t\, \text{div}\, \boldsymbol{a} + \text{div}\,\text{grad}\, \phi)/4\pi \tag{7.3}$$

appears. The relativistic Maxwell equations

$$\text{Grad}_{\text{cov}} F_{\text{con,cov}} = -4\pi J_{\text{cov}} \tag{7.4}$$

now have the following form:

$$\begin{aligned}
&J_{\text{cov}} := (-\sigma^{\text{E}} K^0 + \mathrm{i} j^{\text{E}}_m K^m) + \mathrm{i}(-\sigma^{\text{M}} K^0 + \mathrm{i} j^{\text{M}}_m K^m) \\
&\text{div}\, \boldsymbol{E} = 4\pi\sigma^{\text{E}} \qquad \text{div}\, \boldsymbol{B} = 4\pi\sigma^{\text{M}} \\
&\partial_t \boldsymbol{E} - \text{rot}\, \boldsymbol{B} = -4\pi \boldsymbol{j}^{\text{E}} \qquad -(\partial_t \boldsymbol{B} + \text{rot}\, \boldsymbol{E}) = 4\pi \boldsymbol{j}^{\text{M}}.
\end{aligned} \tag{7.5}$$

Here $\boldsymbol{j}^{\text{E}}$, $\boldsymbol{j}^{\text{M}}$ denote the electric and magnetic current density vectors, respectively. Conservation laws for the electric and magnetic charges can easily be obtained in the same way as for div $\boldsymbol{B} = 0$: the repsective equations of the real and the imaginary components of the scalar part of $\text{Grad}_{\text{con}} J_{\text{cov}}$ are valid for the electric and the magnetic charges in turn. In order to maintain relativistic covariance we can consider the appropriate components of the $F_{\text{cov,con}} J_{\text{cov}}$ product. While the quantities u, $\boldsymbol{S}$ and T remain unchanged, new quantities have to appear. Namely, the force components $\sigma^{\text{M}}\boldsymbol{B} + \boldsymbol{E} \times \boldsymbol{j}^{\text{M}}$ and the energy density $\boldsymbol{B}\boldsymbol{j}^{\text{M}}$ in the field-charge (current) interaction.

8. *Conclusions*

The given representations have the following advantages in contrast to the usual 4×4 matrix description:

(i) Even in the most general case, numerical evaluation of the matrix exponential series can be substituted by computation of simple complex functions as cosh z, sinh z and $z^{1/2}$. This method has similar advantages in relativistic physics as quaternion representation for an arbitrary $\text{O}(3)^+$ rotation in robotics [4, 5].

(ii) The proposed method is proven to be a very concise and powerful algebraic tool to formulate each basic and deduced equation in Maxwell's relativistic electrodynamics.

(iii) It serves as a natural way to formulate relativistic electrodynamics in the case of existence of magnetic monopoles by suggesting appropriate potentials to deduce the magnetic field quantity $\boldsymbol{B}$. Furthermore, it gives expressions of the energy and force density terms to describe the interaction between magnetic monopoles and the electromagnetic field. The formalism makes it also possible to derive the conservation law for the magnetic charge.

The basic disadvantage of the method is that there is no algebraic way to formulate a general symmetric tensor by the use of quaternions. As we have pointed out, it is not a serious obstacle in relativistic electrodynamics.

References

[1] Maxwell J C 1873 *A Treatise on Electricity and Magnetism* (Oxford: Clarendon Press)

[2] Finkelstein D, Jauch J M and Speiser D 1963 Quaternionic representation of compact groups *J. Math. Phys.* **4** (1)

[3] Finkelstein D, Jauch J M and Speiser D 1963 Principle of general Q covarience *J. Math. Phys.* **4** (6)

[4] Chen Z and Hung J C 1987 Application of quaternions in robot control *Proceedings IFAC 10th World Congress on Automatic Control*, vol 4, pp 269–273

[5] Antal K G, Bitó J F and Tar J K 1989 Possibilities of applications of quaternion algebra in robotics *Fizikai Szemle* 412–418 (in Hungarian)

[6] Penrose R 1968 *Structure of Space-Time Battele Recontres 1967 Lectures in Mathematics and Physics* (New York: Benjamin) ch VII

[7] Penrose R and Rindler W 1987 *Two Spinor Calculus and Relativistic Fields, Spinors and Space-Time* vol 1 (Cambridge: Cambridge University Press)

[8] Newman E T and Penrose R 1962 *J. Math. Phys.* **3** 566; 1963 *J. Math. Phys.* **4** 998

[9] Landau L D and Lifshitz E M 1973 *Theoretical Physics II* (Moscow: Nauka)