The Lorentz category in special relativity.

Joachim Lambek
McGill University, Montreal
July 2009.

Abstract

Physical entities may be represented by quaternions with complex components. Conceiving these as arrows in an additive category with two objects, one may encode the information how they transform under a Lorentz transformation, and handle Maxwell's equations and the related Lorentz force. To cope with the Dirac equation we introduce regular matrix representations of the quaternion units and a third object into the category.

1. Introduction.

Special relativity saw the light of day in 1905. Only six years later, two authors independently realized that it could be presented elegantly in terms of *biquaternions*, that is quaternions with complex components. (See Conway [1911, 1912] and Silberstein [1912].)

The idea was to represent physical quantities and operators as biquaternions, together with instructions on how to transform them when the coordinate system was changed by a *Lorentz transformation*, that is a linear transformation which leaves the norm of the biquaternion

$$x = x_0 + i(i_1 x_1 + i_2 x_2 + i_3 x_3)$$

invariant.

A Lorentz transformation takes the form

$$x \mapsto q x q^{*\dagger}$$

where q is a biquaternion of norm $qq^\dagger = q^\dagger q = 1$, $q^\dagger$ being the quaternion conjugate of q. But it turns out that any biquaternion a of physical significance may be transformed as

$$a \mapsto q^X a q^{Y\dagger},$$

where X and Y are superscripts 1 or $*$, $q^1 = q$ and q^* is the complex conjugate of q. (Later we shall also admit the superscript 0 such that $q^0 = 1$, to deal with Dirac spinors.)

Today we may think of the superscripts as objects of an additive category, called a ring with two objects by Barry Mitchell, and known as a *Morita context* in ring theory, and the biquaternions as morphisms. Thus $a : X \to Y$ is to mean that $a \mapsto q^X a q^{Y\dagger}$ when the coordinate system is changed. In particular

$$X \xrightarrow{a} Y \xrightarrow{b} Z$$

means that

$$ab \mapsto q^X a \underbrace{q^{Y\dagger} q^Y} b q^{Z\dagger} = q^X ab q^{Z\dagger}.$$

Note that $a : X \to Y$ implies $a^* : *X \to *Y$, since

$$(q^X a q^{Y\dagger})^* = q^{*X} a^* q^{*Y\dagger},$$

but $a^\dagger : Y \to X$, since

$$(q^X a q^{Y\dagger})^\dagger = q^Y a^\dagger q^{X\dagger}.$$

(Later we will introduce a third object 0 to help representing spinors.)

2. Some Hermitian biquaternions in physics.

Of special interest are *Hermitian* biquaternions a for which $a^* = a^\dagger$, i.e. $a^{*\dagger} = a$, in other words

$$a = a_0 + i(i_1 a_1 + i_2 a_2 + i_3 a_3),$$

when $a_0,\ a_1,\ a_2$ and a_3 are real and i_1, i_2 and i_3 such that

$$i_1^2 = i_2^2 = i_3^2 = i_1 i_2 i_3 = -1$$

are the basic quaternion units.

In particular, the following Hermitian biquaternions played a rôle in the early development:

space-time:

$$x = x_0 + i(i_1 x_1 + i_2 x_2 + i_3 x_3) = t + i\mathbf{x},$$

where $x_0 = t$ represents time and $\mathbf{x}$ position in 3-space;

energy-momentum:

$$p = p_0 + i(i_1 p_1 + i_2 p_2 + i_3 p_3) = \varepsilon + i\mathbf{p},$$

where $p_0 = \varepsilon$ represents energy and $\mathbf{p}$ momentum;

four-potential:

$$A = A(x) = \varphi + i(i_1 A_1 + i_2 A_2 + i_3 A_3) = \varphi + i\mathbf{A},$$

where $e\varphi$ represents the potential energy and $e\mathbf{A}$ the potential momentum, e being the electric charge of a moving particle;

partial differentiation:

$$D = \frac{\partial}{\partial x_0} - i\left(i_1 \frac{\partial}{\partial x_1} + i_2 \frac{\partial}{\partial x_2} + i_3 \frac{\partial}{\partial x_3}\right) = \frac{\partial}{\partial t} - i\nabla.$$

All of these are to be thought of as arrows $1 \to *$, thus transforming like $x \mapsto q x q^{*\dagger}$, hence possessing an invariant norm. For example, the Hermitian differential dx has norm

$$ds^2 = dx\,dx^\dagger = dx^\dagger dx,$$

where s is called the *interval*, and the energy-momentum p has norm

$$m^2 = pp^\dagger = p^\dagger p,$$

where m is known as the *rest-mass*.

When $m \neq 0$ one requires that also $ds \neq 0$ and writes

$$\begin{aligned} p = \quad & m\frac{dx}{ds} = m\frac{dt}{ds}\left(1 + i\frac{d\mathbf{x}}{dt}\right) \\ = \quad & \quad m(1-v^2)^{-\frac{1}{2}}(1+i\mathbf{v}), \end{aligned}$$

where

$$\mathbf{v} = i_1\frac{dx_1}{dt} + i_2\frac{dx_2}{dt} + i_3\frac{dx_3}{dt}$$

is the classical velocity vector of norm v^2. (We have chosen units so that the velocity of light and the dielectric constant are 1 and Planck's constant is 2π.)

In the absence of an external force, one postulates *conservation* of energy-momentum $\frac{dp}{ds} = 0$. In the presence of an electro-magnetic force, the so-called *Lorentz force*, one would expect

$$\frac{d}{ds}(p + eA) = 0,$$

A being the electro-magnetic four-potential, hence eA the energy-momentum.

It used to be assumed that there is no way of measuring $\mathbf{A}$, but that $A = \varphi + i\mathbf{A}$ is only known by the vector part of D^*A:

$$V(D^*A) = -(\mathbf{B} + i\mathbf{E}),$$

where $\mathbf{B}$ and $\mathbf{E}$ are the magnetic and electric fields respectively. But this is the same as the vector part of $D^*(A - D\theta)$, where $\theta = \theta(x)$ is a scalar, since $D^*D\theta$ is a scalar. Thus, the measurable effect of A appeared to be invariant under the *gauge transformation*

$$A \mapsto A_\theta = A - D\theta.$$

This was traditionally exploited by assuming that the scalar part $D \odot A = 0$. We will refrain from making this assumption here. The gauge transformation also does not affect Maxwell's equation, which asserts that

$$D(\mathbf{B} + i\mathbf{E}) = DD^*A = -J,$$

where J denotes the *charge-current density.*

3. Working with Hermitian biquaternions.

Recalling that $a : X \to Y$ implies $a^* : X* \to Y*$ and $a^\dagger : Y \to X$, we infer that it implies $a^{*\dagger} : Y* \to X*$. In particular, if $a : 1 \to *$, then

$$a^{*\dagger} : 1 = ** \to *,$$

hence also

$$H(a) = \frac{1}{2}(a + a^{*\dagger}),\ a - H(a) = \frac{1}{2}(a - a^{*\dagger}) : 1 \to *.$$

We call $H(a)$ the *Hermitian part* of a and $a - H(a) = -iH(ia)$ the *skew-Hermitian part.*

On the other hand, if $a : 1 \to 1$ the $a^\dagger : 1 \to 1$, hence also

$$S(a) = \frac{1}{2}(a + a^\dagger),\ V(a) = \frac{1}{2}(a - a^\dagger) : 1 \to 1.$$

We call these the *scalar* and *vector* parts of a respectively.

From the Hermitian biquaternions $a, b : 1 \to *$ we can construct the composite

$$1 \xrightarrow{a} * \xrightarrow{b^\dagger} 1,$$

which we will write as

$$ab^\dagger : 1 \to 1,$$

the order of multiplication being opposite from the usual categorical convention.

If $a = a_0 + i\mathbf{a}$ and $b = b_0 + i\mathbf{b}$, we have

$$ab^\dagger = a_0 b_0 - \mathbf{a} \circ \mathbf{b} + \mathbf{a} \times \mathbf{b} + i(\mathbf{a} b_0 - a_0 \mathbf{b}),$$

where the scalar part

$$a \odot b = S(ab^\dagger) = a_0 b_0 - \mathbf{a} \circ \mathbf{b}$$

and the vector part

$$ab - a \odot b = V(ab^\dagger) = \mathbf{a} \times \mathbf{b} + i(\mathbf{a} b_0 - a_0 \mathbf{b})$$

both represent arrows $1 \to 1$. Note that $a \odot b = b \odot a$, but $V(ab^\dagger) = -V(ba^\dagger)$.

If a, b, c are three Hermitian biquaternions $* \to 1$, we have

$$* \xrightarrow{a} 1 \xrightarrow{b^*} * \xrightarrow{c} 1,$$

so that

$$ab^*c : \; * \to 1.$$

A routine calculation shows that its Hermitian part

$$H(ab^*c) = a(b \odot c) - b(c \odot a) + c(a \odot b). \tag{1}$$

This identity first appears in Gürsey [1955]. We will apply it in case $a = e\dfrac{dx}{ds}$, $b = D$ and $c = A$.

4. Conservation of energy-momentum.

From (1) we obtain

$$-H\left(e\frac{dx}{ds}D^*A\right) + e\frac{dx}{ds}(D \odot A) - D\left(e\frac{dx}{ds} \odot A\right) + e\left(\frac{dx}{ds} \odot D\right)A = 0. \tag{2}$$

Here the first two terms combine to

$$\begin{aligned}
- \quad & H\left(e\frac{dx}{ds}(D^*A - D \odot A)\right) \\
= \quad & H\left(e\frac{dx}{ds}(\mathbf{B} + i\mathbf{E})\right) \\
= \quad & e\frac{d\mathbf{x}}{ds} \circ \mathbf{E} + ie\left(\frac{d\mathbf{x}}{ds} \times \mathbf{B} + \frac{dt}{ds}\mathbf{E}\right) \\
= \quad & e\frac{dt}{ds}\left(\mathbf{v} \circ \mathbf{E} + i(\mathbf{v} \times \mathbf{B} + \mathbf{E})\right).
\end{aligned}$$

We recognize $e(\mathbf{v}\times\mathbf{B}+\mathbf{E})$ as the *classical Lorentz force* $\dfrac{d\mathbf{p}}{dt}$, so we must interpret $e(\mathbf{v}\circ\mathbf{E})$ as the *classical Lorentz power* $\dfrac{d\varepsilon}{dt}$, and the whole expression as a relativistic version of the Lorentz force $\dfrac{dp}{ds}$.

Thus (2) asserts that

$$\frac{d}{ds}(p+eA)=\frac{dp}{ds}+\frac{deA}{ds}=D\left(A\odot e\frac{dx}{ds}\right),$$

where $p+eA$ is the kinetic plus potential energy-momentum. At first sight, the right hand side appears puzzling, but it seems to be related to the quantum-mechanical *Bohm-Aharonov* effect. (See Ryder [1996]). Anyway, if we write

$$A\odot\frac{dx}{ds}=\frac{d\theta}{ds},\ \ A_\theta=A-D\theta,$$

we have

$$\frac{d}{ds}(p+eA_\theta)=0$$

and we may as well redefine the 4-potential as A_θ, thus establishing that the total energy momentum is conserved.

There is no point in repeating this process, since

$$\begin{aligned}\frac{dx}{ds}\odot A_\theta &= \frac{dx}{ds}\odot A-\frac{dx}{ds}\odot D\theta\\ &= \frac{dx}{ds}\odot A-\frac{d\theta}{ds}=0.\end{aligned}$$

Hence we may as well assume that $A=A_\theta$ in the first place, i.e. $\dot{x}\odot A=0$, i.e. the potential is "orthogonal" to the current in Minkowski space.

In summary of the above argument, Maxwell's equations show how the electro-magnetic field depends on the charge-current density and may be derived from a four-potential, which is determined only up to a gauge transformation. The path of a charged particle is influenced by the Lorentz force due to this field. Aided by a suitable gauge transformation, we may assume that the underlying four-potential is orthogonal to the path of the given particle in Minkowski space. When multiplied by the charge, the four-potential then gives rise to the potential energy-momentum of the charged particle, and the relativistic Lorentz force may be accounted for by the conservation of its total (kinetic plus potential) energy-momentum. Physicists must be familiar with this observation, though I have not seen it mentioned in any of the texts I have consulted.

5. Dirac spinors.

In vacuum, the photon, or any particle of mass zero, satisfies the equation $DD^*A=0$. This has been generalized to

$$DD^*A+m^2A=0$$

for particles of mass $m \neq 0$, the so-called *Klein-Gordon* equation. This second order differential equation may be replaced by two first order equations, provided $m \neq 0$:

(3) $$D^*A = mA',\ DA' = -mA.$$

For bosons of spin 1,

$$A : 1 \longrightarrow *,\ A' : * \longrightarrow *$$

but for fermions of spin $\frac{1}{2}$ we postulate

$$A : 1 \longrightarrow 0,\ A' : * \longrightarrow 0$$

and then A and A' are called *spinors*. We have now introduced a third object 0 into our category.

The two equations of (3) may be combined into one:

$$D^*\Psi = -mj\Psi,$$

the so-called Dirac equation, where

$$\Psi = A + jA' :\ 1 \longrightarrow 0$$

provided

$$j^2 = -1,\ ji = -ij.$$

This looks odd when i is the usual complex number, but becomes realistic if left multiplication by i and j is replaced by right multiplication with the basic quaternions i_1 and i_2 respectively.

(Note that $q^0 = 1$, hence the set $\{0, *, 1\}$ behaves like $\mathbb{Z}$ mod 3 under multiplication. The arrow $\Psi : 1 \longrightarrow 0$ represents the fact that $\Psi \mapsto q\Psi$ under a Lorentz transformation.)

More precisely, we identify the quaternion units i_α ($\alpha = 0$ to 3) with their left regular matrix representation and write j_α for the corresponding right representation.

The Dirac equation then takes the form

$$(D^* + mj_2)\Psi = 0,$$

where $D^* = \dfrac{\partial}{\partial t} + j_1\nabla$. (There is no need for introducing the ad hoc Pauli matrices.)

Since the Klein-Gordon equation has the explicit solution

$$A = \cos(x \odot p)\Psi_0,$$

Ψ_0 being a constant matrix, our procedure yields a solution to the Dirac equation:

$$\Psi = (\cos(x \odot p) + \eta \sin(x \odot p))\Psi_0,$$

where

$$-\eta = j_2 dx^\dagger/ds = (dx/ds)j_2.$$

Taking advantage of the fact that $\eta^2 = -1$, we may write this more elegantly thus:

$$\Psi = \exp(\eta(x \odot p))\Psi_0.$$

In the Dirac equation, Ψ may be multiplied by any constant matrix on the right, in particular by a column vector $[c]$, where c is a quaternion with real components c_α, and $[c]$ is the transposed of the row vector (c_0, c_1, c_2, c_3), rendering $\Psi[c]$ into a column vector itself, say of the form $[\psi]$, ψ being a real quaternion. Then

$$i_\alpha[\psi] = [i_\alpha\psi],\ j_\alpha[\psi] = [\psi i_\alpha],$$

hence the Dirac equation with $\Psi = [\psi]$ takes the quaternionic form

$$\frac{\partial\psi}{dt} + \nabla\psi i_1 + m\psi i_2 = 0.$$

We have taken j_α to be the right regular representation of i_α, but we might as well have taken it to be that of $ri_\alpha r^\dagger$, where r is any quaternion of norm 1. Then the above equation takes the form

$$\frac{\partial\psi}{\partial t}r + \nabla\psi r i_1 + m\psi r i_2 = 0,$$

which amounts to replacing ψ by ψr.

REFERENCES

A.W. Conway, On the applications of quaternions to some recent development of electrical theory, *Proc. Royal Inst. Academy* **A29**(1911), 80.

A.W. Conway, The quaternion form of relativity, *Phil. Mag.* **24**(1912), 208.

A.W. Conway, Quaternions and quantum mechanics, *Pontificia Academia Scientiarum* **12**(1948), 204-277.

R.P. Feynman, R.B. Leighton and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley Publ. Co., Reading Mass 1963/1977.

F. Gürsey, Contributions to the quaternionic formalism in special relativity, *Rev. Faculté Sci.* Univ. Istanbul A20(1955), 149-171.

J. Lambek, If Hamilton had prevailed: Quaternions in Physics, *The Math. Intelligencer* **17**(1995), 7-15.

J. Lambek, Quaternions and the particles of nature, *Report* 96-03, Department of Mathematics, McGill University, 1996.

R. Penrose, *The road to reality: a complete guide to the laws of the universe*, Jonathan Cape, London 2004.

L.H. Ryder, Quantum field theory, Cambridge University Press, 1985/1996.

L. Silberstein, Quaternionic form of relativity, *Phil. Mag.* **23**(1912), 790.

L. Silberstein, *Theory of relativity*, London, McMillan (1924).

A. Sudbery, *Quantum mechanics and the particles of nature*, Cambridge University Press, 1986.