A quaternion formulation of the Dirac equation

Maurits H. Silvis [*]
Advisor: Mees de Roo [‡]

July 11, 2010

Abstract

In this thesis the formalisms of quaternions and biquaternions have been employed to reformulate Dirac's relativistic wave equation and to investigate claims concerning elegance, intuitiveness and new physical results of such a formulation. In this fashion, an elegant formulation of the Dirac equation in terms of biquaternions was found. After this, reproduction of some well-known physical results like plane wave solutions and the nonrelativistic approximation were achieved. It must be admitted, however, that there seems to be no reason to believe that a quaternion or biquaternion formulation of the Dirac equation contains any additional physics and the purpose of getting used to a different formalism than the usual may be questioned as it does not seem to have deepened the understanding of the Dirac equation.

[*] University of Groningen; student no.: 1691589; e-mail: m.h.silvis@student.rug.nl
[‡] Centre for Theoretical Physics, University of Groningen

Contents

Chapter 1

Introduction

In 1928 Paul Dirac formulated a first-order relativistic wave equation, now called the Dirac equation, thereby postulating the existence of antiparticles. Space and time appear on an equal footing in this equation like in all of special relativity. Due to the apparent four-dimensionality of our world that follows from this it has been suggested that a formalism using quaternions would be natural for all of relativistic quantum mechanics.

The quaternions form an extension of the complex numbers invented by William Rowan Hamilton in 1843 and were introduced in his *Lecture on quaternions* [1]. They have four components and were designed to describe rotations in three-dimensional space and, thus, they can be seen as a generalization of the complex numbers – having two components – that so elegantly describe rotations in the plane.

Whether elegant or not, at the end of the nineteenth century the vector analysis developed by Heaviside and Gibbs marked the demise of quaternions [2]. Although quaternions never again entered mainstream science, the formalism has had its own group of devotees.

As early as 1912, Arthur Conway [3] was interested in the application of complex quaternions or biquaternions to the special theory of relativity. Later, he used biquaternions to find a simpler formulation of the Dirac equation, but not before Silberstein and Lanczos tried to apply quaternions to physics [4]. In fact, even Dirac [5] himself set out to investigate the applications of quaternions in physics and successfully finds a relation between Lorentz transformations and quaternions.

In more recent years, others have attempted to revive interest in the quaternion formalism by trying to formulate and investigate complex quaternion formulations of the Dirac equations [6, 7, 8]. Also, the connection with Lorentz transformations and the very elegant description of Maxwell's equations using biquaternions has been studied repeatedly [9, 10]. Joachim Lambek [11] gives a nice account of this, also treating a biquaternion version of the Dirac equation for the electron in an electromagnetic field.

Clearly, the attempts of these authors to give quaternions a place in physics have failed. Suggestions of elegance of the quaternion formalism and the power of quaternions in providing us with intuitive descriptions of physics are to be found in almost all of the aforementioned references, however. Some articles even contain claims of finding "new physics" behind a different formulation of the Dirac equation [12]. Patrick Girard [13] seems to imply we have to take these suggestions seriously, for he points out that the groups of normal and complex quaternions can be related to many important groups in physics, among which are the special unitary group SU(2) and the Lorentz group.

In this thesis, our aim will be to investigate the elegance, intuitiveness and power of a quaternion formalism in physics by trying to find a (complex) quaternion formulation of the Dirac equation and interpreting the physical results it implies.

To this end, we will lay down the basis of the algebra of quaternions and complex quaternions and provide a link of the latter with four-vectors in the upcoming chapter. After this exploration of the properties of quaternions, we will look at the Klein-Gordon and Dirac equations in chapter 3. Also, (complex) quaternion representations of the latter of these equations are investigated. This naturally leads us to the following chapter that treats (bi)quaternion formulations of the Dirac equation. The physical content of this equation is investigated in chapter 6, after devoting a chapter to biquaternion formulations of electromagnetism. In the final chapter, a comprehensive account of our conclusions is given.

Chapter 2

The algebra of quaternions

2.1 Introduction

Before we start to speak of the Dirac equation and other physical concepts, it will be useful to explore the properties of quaternions and biquaternions. In the next section we will investigate the multiplication of quaternions and find out what representations can be used to describe them. Also, we will try to relate quaternions with four-vectors, making our first move towards physics. As some problems are encountered in this identification, we will extend our quaternion algebra to a biquaternion algebra. Section 2.3 will be devoted to this, where we will speak of the multiplicative properties and representations of biquaternions, after which we will relate them to four-vectors. In the final section we will summarize the most important findings of this chapter.

2.2 Quaternions

2.2.1 Properties of quaternions

The *quaternions* form an extension of the complex numbers in such a way that besides the usual complex unit i two more are introduced, namely j and k. Hence, when we write a complex number z as

$$z = a^0 + a^1 i \tag{2.2.1}$$

a quaternion q would look like

$$q = a^0 + a^1 i + a^2 j + a^3 k \tag{2.2.2}$$

Here, the parameters a^0, a^1, a^2 and a^3 are real numbers. The superscript numbers are not powers of a, but so called *upper indices* used to distinguish the four coefficients of the quaternion. We will write $q \in \mathbb{H}$ to signify that q is a quaternion. The real number a^0 is generally referred to as the *scalar part* of q. When one speaks of the *vector part* of a quaternion, $a^1 i + a^2 j + a^3 k$ is meant, which is equivalent to a vector with three real components. Therefore, sometimes we will write

$$q = a^0 + \mathbf{a} \tag{2.2.3}$$

A quaternion that only has a scalar part is a *real number*. Quaternions with only a vector part are called *pure imaginary*.

The quaternion units i, j and k, that all commute with 1, satisfy a number of interesting properties:

$$i^2 = j^2 = k^2 = ijk = -1 \tag{2.2.4}$$

From these properties we can deduce that quaternion multiplication is not commutative, that is, in general $qr \neq rq$ for two quaternions q and r and that due to this anticommutativity multiplication of the quaternion units is similar to the cross product in three dimensions. Consider for instance,

$$\left.\begin{array}{l} (ijk) * k = ijk^2 = -ij \\ -1 * k = -k \end{array}\right\} \qquad ij = k \tag{2.2.5}$$

whereas

$$\left.\begin{array}{l} ji * (ijk) = k \\ ji * -1 = -ji \end{array}\right\} \quad -ji = k \tag{2.2.6}$$

The resulting multiplication table of the basis elements 1, i, j and k looks as given in Table 2.1.

Table 2.1: Quaternion multiplication table

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Mathematically speaking, the basis elements plus their negative counterparts form a *group* of order eight, that is, a group of eight elements. Usually this group is denoted Q_8.

From the multiplication table given one can obtain the general product of two quaternions q and r

$$\begin{aligned} qr = & (a^0 + \mathbf{a})(b^0 + \mathbf{b}) = (a^0 + a^1 i + a^2 j + a^3 k)(b^0 + b^1 i + b^2 j + b^3 k) \\ = & a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 + \\ & (a^0 b^1 + a^1 b^0 + a^2 b^3 - a^3 b^2) i + \\ & (a^0 b^2 + a^2 b^0 + a^3 b^1 - a^1 b^3) j + \\ & (a^0 b^3 + a^3 b^0 + a^1 b^2 - a^2 b^1) k \\ = & a^0 b^0 - \mathbf{a} \cdot \mathbf{b} + a^0 \mathbf{b} + \mathbf{a} b^0 + \mathbf{a} \times \mathbf{b} \end{aligned} \tag{2.2.7}$$

As can be seen from the above result the multiplication of two vector parts produces the following peculiar but handy result, having no direct analogy in vector analysis. For clarity: $\mathbf{a} \cdot \mathbf{b}$ signifies the usual dot product of two vectors $\mathbf{a}$ and $\mathbf{b}$, $a^1 b^1 + a^2 b^2 + a^3 b^3$.

$$\mathbf{ab} = -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b} \tag{2.2.8}$$

Analogous to the complex conjugate of our complex number z

$$z^* = a^0 - a^1 i \tag{2.2.9}$$

we introduce *quaternion conjugation* denoted by † for which

$$1^\dagger = 1, i^\dagger = -i, j^\dagger = -j \text{ and } k^\dagger = -k \tag{2.2.10}$$

resulting in the quaternion conjugate of q

$$q^\dagger = a^0 - a^1 i - a^2 j - a^3 k \tag{2.2.11}$$

Conjugation allows us to distinguish the aforementioned special types of quaternions: real numbers and pure imaginary quaternions. When $q^\dagger = q$ only the coefficient a^0 is nonzero and thus q is a real number. When $q^\dagger = -q$, q is pure imaginary.

Furthermore, using conjugation we can define the (squared) *modulus* (or *norm* or *magnitude*) of a quaternion to be

$$\begin{aligned} |q|^2 := & q^\dagger q = q q^\dagger = (a^0 + a^1 i + a^2 j + a^3 k)(a^0 - a^1 i - a^2 j - a^3 k) \\ = & (a^0)^2 - a^0 a^1 i - a^0 a^2 j - a^0 a^3 k + \\ & (a^1)^2 + a^0 a^1 i + a^1 a^3 j - a^1 a^2 k + \\ & (a^2)^2 - a^2 a^3 i + a^0 a^2 j + a^1 a^2 k + \\ & (a^3)^2 + a^2 a^3 i - a^1 a^3 j + a^0 a^3 k \\ = & (a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2 = (a^0)^2 + \mathbf{a} \cdot \mathbf{a} \end{aligned} \tag{2.2.12}$$

Note that we could also have obtained this result by taking $q = q$ and $r = q^\dagger$ in the general product (2.2.7).

$$\begin{aligned} qq^\dagger &= (a^0 + \mathbf{a})(a^0 - \mathbf{a}) = (a^0)^2 - \mathbf{a}\cdot -\mathbf{a} - a^0\mathbf{a} + \mathbf{a}a^0 + \mathbf{a}\times -\mathbf{a} \\ &= (a^0)^2 + \mathbf{a}\cdot\mathbf{a} \end{aligned} \tag{2.2.13}$$

2.2.2 Representations of quaternions

With Table 2.1, the multiplication table of the quaternion units, we can derive what the conjugacy classes of the quaternion group are. As both 1 and -1 commute with the other elements from the group, they must form separate conjugacy classes. It turns out that the elements i and $-i$ are in the same conjugacy class and from symmetry then follows that $\pm j$ and $\pm k$ form conjugacy classes as well.

Using these classes we can construct Table 2.2, the character table of the quaternion group.

Table 2.2: Character table of the quaternion group

$\chi(K_l)$	K_1 (1)	K_2 (-1)	K_3 ($i,-i$)	K_4 ($j,-j$)	K_5 ($k,-k$)
$D^{(1)}$	1	1	1	1	1
$D^{(2)}$	1	1	1	-1	-1
$D^{(3)}$	1	1	-1	1	-1
$D^{(4)}$	1	1	-1	-1	1
$D^{(5)}$	2	-2	0	0	0

Note that this character table is the same as that for the group D_4, the dihedral group of order eight, although the quaternion group Q_8 and D_4 are not isomorphic [14].

Complex representation of quaternions

We are interested in representations of dimension two and more and therefore we can in fact discard any row of the character table but the last one. This last row requires that our unit 1 should be represented by a 2×2 matrix with trace 2. It is exactly the unit or identity matrix, 1_2, that satisfies this property. So, in writing D instead of $D^{(5)}$,

$$1 \longleftrightarrow D(1) = 1_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.2.14}$$

There are three conjugacy classes that are to be represented by a 2×2 matrix with trace 0. This suggests we can invoke the Pauli matrices, which will be treated in more detail in section 3.4,

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.2.15}$$

that satisfy

$$\sigma^l\sigma^m = \delta^{lm}1_2 + i\epsilon^{lmn}\sigma^n, \qquad \text{for } l, m = 1, 2, 3 \tag{2.2.16}$$

Although it may seem a convenient choice to relate i, j and k to σ^1, σ^2 and σ^3 respectively this will not work as the above property is not satisfied for all values of l and m in this way. However, when we identify i with $\sigma^1\sigma^2 = i\sigma^3$, j with $\sigma^3\sigma^2 = i\sigma^2$ and k with $\sigma^2\sigma^3 = i\sigma^1$ we get a correct way of representing the quaternion units.

$$i \longleftrightarrow i\sigma^3 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, j \longleftrightarrow i\sigma^2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, k \longleftrightarrow i\sigma^1 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \tag{2.2.17}$$

This provides us with the following irreducible 2×2 complex matrix representation of a general quaternion q.

$$q = a^0 + a^1 i + a^2 j + a^3 k \longleftrightarrow D(q) = \begin{bmatrix} a^0 + a^1 i & a^2 + a^3 i \\ -a^2 + a^3 i & a^0 - a^1 i \end{bmatrix} \tag{2.2.18}$$

Now it may become clear why we chose † as the symbol for quaternion conjugation. In quantum mechanics this symbol is widely used to denote the *Hermitian conjugate* or complex conjugate of a transposed matrix. It happens to be the case that in our representation (2.2.18) taking the Hermitian conjugate exactly matches quaternion conjugation.

$$
\begin{aligned}
q^\dagger = a^0 - a^1 i - a^2 j - a^3 k \longleftrightarrow D(q^\dagger) &= \begin{bmatrix} a^0 - a^1 i & -a^2 - a^3 i \\ a^2 - a^3 i & a^0 + a^1 i \end{bmatrix} \\
&= \begin{bmatrix} a^0 + a^1 i & -a^2 + a^3 i \\ a^2 + a^3 i & a^0 - a^1 i \end{bmatrix}^* = \begin{bmatrix} a^0 + a^1 i & a^2 + a^3 i \\ -a^2 + a^3 i & a^0 - a^1 i \end{bmatrix}^\dagger = D(q)^\dagger
\end{aligned} \tag{2.2.19}
$$

Real representation of quaternions

When we take the stated irreducible 2×2 complex representation (2.2.17) and make the replacements

$$
1 \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, 0 \rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, i \rightarrow \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{2.2.20}
$$

we get a reducible 4×4 real representation that is a direct sum of two of the previous 2×2 complex representations. We now have

$$
q = a^0 + a^1 i + a^2 j + a^3 k \longleftrightarrow D(q) = \begin{bmatrix} a^0 & a^1 & a^2 & a^3 \\ -a^1 & a^0 & -a^3 & a^2 \\ -a^2 & a^3 & a^0 & -a^1 \\ -a^3 & -a^2 & a^1 & a^0 \end{bmatrix} \tag{2.2.21}
$$

A feature of this representation is that a pure imaginary quaternion, that is a quaternion with $a^0 = 0$ is represented by an antisymmetric matrix. This may prove useful in finding a quaternion representation of electrodynamics.

2.2.3 Identification of four-vectors and tensors with quaternions

Four-vectors

Quaternions and four-vectors both have four independent components and this implies we can use the following identification between them. For clarity we have given the quaternion the name x and its real parameters the names x^0 to x^3.

$$
x = x^0 + x^1 i + x^2 j + x^3 k \longleftrightarrow x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{2.2.22}
$$

Introducing the *Minkowski metric*, a matrix usually denoted with the Greek letter η

$$
\eta = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix} \tag{2.2.23}
$$

we can easily identify quaternion conjugation with application of this matrix, signifying the change from a contravariant to a covariant vector.

$$
x^\dagger = x^0 - x^1 i - x^2 j - x^3 k \longleftrightarrow x_\mu = \eta_{\mu\nu} x^\nu = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{bmatrix} \tag{2.2.24}
$$

It is very important to observe here that only in the four-vector formalism upper and lower indices have a meaning in the sense of contravariance and covariance. In the quaternion formalism this difference is signified by the quaternion units having a plus or minus sign and thus the quaternion coefficients could just as well have been written x_0 to x_3 here, or even without an index, like a, b, c and d.

Continuing with our identification of four-vectors with quaternions, we observe something does not work out. The squared norm of a four-vector is defined to be

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = \eta_{\mu\nu} x^\mu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x^0)^2 - \mathbf{x}\cdot\mathbf{x} \tag{2.2.25}$$

whereas we found the squared norm of a quaternion in (2.2.12) as

$$|x|^2 := x^\dagger x = (x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 = (x^0)^2 + \mathbf{x}\cdot\mathbf{x} \tag{2.2.26}$$

and this last expression will not be invariant under Lorentz transformations.

Note that in (2.2.25) we have used the Einstein summation convention, that is, summation is implicitly assumed whenever we encounter an index twice, once in an upper and once in a lower position, like μ in this case. Generally, μ, ν and ρ are used as indices that run from 0 to 3.

Tensors

At the end of section 2.2.2 we remarked that the possibility to represent the vector part of a quaternion by an antisymmetric matrix could provide us with a representation of classical electromagnetism. Let us explain briefly what we mean by this, in anticipation of chapter 5. Looking at the second-rank tensor containing the electromagnetic field strengths there, we observe it is antisymmetric, just like our representation is. Despite this promising point, a problem can be recognized straight-away. The antisymmetric tensor contains six independent components, so it cannot be identified with the vector part of a single quaternion that consists of three numbers. Perhaps, identification with multiplication from both left and right with pure imaginary quaternions solves this problem, but we will not look into this.

2.2.4 Towards a biquaternion algebra

As was just noted, some problems arise when we try to relate four-vectors and tensors to quaternions. We will see that a feasible solution to this problem is to extend the quaternions to the biquaternions, which we will in the next section. Before we go over to this approach, let us explain why not everyone is fond of this solution, however.

First of all, the quaternion group takes a special place in mathematics as it is one of only three groups, the others being the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$, that form a division ring. This means that $q \neq 0$ and $r \neq 0$ imply $qr \neq 0$ for any two quaternions q and r. The group of complex quaternions is not as special as this, seen from the mathematical point of view. This group is no division ring and hence it is possible to multiply two nonzero biquaternions and still end up with the remarkable result of zero. For instance,

$$(1 - ii_1)(1 + ii_1) = 1 - (ii_1)^2 = 0 \tag{2.2.27}$$

Secondly, the biquaternions have eight components instead of four as the quaternions do. Dirac remarks that this seems to be too many for a simple description of physics and this had led him to focus on normal quaternions. He finds a direct relation between Lorentz and quaternion transformations in his 1945 paper. [5]

Furthermore, in the quaternion formalism there are workarounds for the problem concerning the Lorentz invariant encountered in the previous section. As one can see from the general product of two quaternions, the scalar part a quaternion squared provides a candidate for Lorentz invariance as

$$\begin{aligned} xx &= (x^0 + x^1 i + x^2 j + x^3 k)^2 \\ &= x^0x^0 - x^1x^1 - x^2x^2 - x^3x^3 + (\ldots)i + (\ldots)j + (\ldots)k \end{aligned} \tag{2.2.28}$$

This method is suggested by Stefano de Leo in a paper where he also succeeds in finding the generators of the Lorentz group in terms of normal quaternions. [9]

However a very interesting direction of research, we will not pursue a quaternion identification with four-vectors, because we will see in section 2.3.4 that a different formalism provides us with a more straight-forward Lorentz invariant than (2.2.28) does and we will focus on the Dirac equation – not on Lorentz transformations.

2.3 Biquaternions

2.3.1 Properties of biquaternions

In the previous section, we extended the plane of the complex numbers to the four-dimensional space of the quaternions. Now, we will extend the quaternions to the *biquaternions* or *complex quaternions*. This last term is suggestive of how this extensions works: we will use complex coefficients now instead of the real ones used before. Therefore, a biquaternion comes from the product space $\mathbb{C} \otimes \mathbb{H}$.

A problem arises here that the reader may be aware of already. We need to distinguish between the complex unit i in the complex parameters and the quaternion unit i, for otherwise we simply end up with a normal quaternion. That is, we require a square root of -1 independent of i, j and k.

To this end we will write the quaternion units i, j and k as i_1, i_2 and i_3 whenever we speak of biquaternions. These (new) units are defined to commute with the complex unit of the biquaternion coefficients, which we will still denote with i, thus

$$[i_l, i] = 0 \quad \text{for } l = 1,2,3 \tag{2.3.1}$$

A biquaternion q may now be written

$$q = c^0 + c^1 i_1 + c^2 i_2 + c^3 i_3 = c^0 + c^l i_l = c^0 + \mathbf{c} \tag{2.3.2}$$

where c^0, c^1, c^2 and c^3 are complex numbers. Here, we have used the summation convention again. For biquaternions the doubly occurring index – l, m or n – will always run from 1 to 3 to distinguish complex quaternions from four-vectors.

Note that the multiplicative properties of the quaternion units, (2.2.4), can easily be summarized using our current notation.

$$i_l i_m = -\delta_{lm} + \epsilon_{lmn} i_n \quad \text{for } l, m = 1,2,3 \tag{2.3.3}$$

Due to the fact that we now have a complex unit besides the quaternion units, the biquaternion group consists of sixteen elements and its multiplication table looks as given in Table 2.3.

Table 2.3: Biquaternion multiplication table

	1	i_1	i_2	i_3	i	ii_1	ii_2	ii_3
1	1	i_1	i_2	i_3	i	ii_1	ii_2	ii_3
i_1	i_1	-1	i_3	$-i_2$	ii_1	$-i$	ii_3	$-ii_2$
i_2	i_2	$-i_3$	-1	i_1	ii_2	$-ii_3$	$-i$	ii_1
i_3	i_3	i_2	$-i_1$	-1	ii_3	ii_2	$-ii_1$	$-i$
i	i	ii_1	ii_2	ii_3	-1	$-i_1$	$-i_2$	$-i_3$
ii_1	ii_1	$-i$	ii_3	$-ii_2$	$-i_1$	1	$-i_3$	i_2
ii_2	ii_2	$-ii_3$	$-i$	ii_1	$-i_2$	i_3	1	$-i_1$
ii_3	ii_3	ii_2	$-ii_1$	$-i$	$-i_3$	$-i_2$	i_1	1

Now, using the given multiplication table, we see that the general product of two biquaternions q and r looks the same as the general product of two quaternions (2.2.7).

$$\begin{aligned} qr = &\ (c^0 + \mathbf{c})(d^0 + \mathbf{d}) \\ = &\ c^0 d^0 - \mathbf{c} \cdot \mathbf{d} + c^0 \mathbf{d} + \mathbf{c} d^0 + \mathbf{c} \times \mathbf{d} \end{aligned} \tag{2.3.4}$$

There is a difference, however. The numbers c^0, c^1, $\ldots d^3$ are complex numbers here, unlike the real numbers a^0, a^1, $\ldots$, b^3. As we will see, after speaking of conjugation of a biquaternion this results in some interesting possibilities.

The introduction of a complex unit besides the quaternion units not only results in a large multiplication table, it also allows for different types of conjugation.

The quaternion conjugate of a biquaternion produces a result that is very similar to the conjugate of a normal quaternion (2.2.11), because we have

$$1^\dagger = 1, \quad i_l^\dagger = -i_l \quad \text{for } l = 1,2,3 \tag{2.3.5}$$

resulting in the quaternion conjugate of the complex quaternion q,

$$q^\dagger = c^0 - c^1 i_1 - c^2 i_2 - c^3 i_3 = c^0 - c^l i_l = c^0 - \mathbf{c} \tag{2.3.6}$$

As c^0 to c^3 are complex numbers, we could also take the complex conjugate of q,

$$q^* = (c^0)^* + (c^1)^* i_1 + (c^2)^* i_2 + (c^3)^* i_3 = (c^0)^* + (c^l)^* i_l = (c^0)^* + \mathbf{c}^* \tag{2.3.7}$$

or a combination of both

$$q^{*\dagger} = (c^0)^* - (c^1)^* i_1 - (c^2)^* i_2 - (c^3)^* i_3 = (c^0)^* - (c^l)^* i_l = (c^0)^* - \mathbf{c}^* \tag{2.3.8}$$

These different conjugation modes allow us to distinghuish several types of biquaternions. We will look at them in section 2.3.3, just as soon as we have described the representations of complex quaternions.

2.3.2 Representations of biquaternions

In the biquaternion group there are four elements that commute with all others, namely ± 1 and $\pm i$ and hence, they must form separate conjugacy classes. In the quaternion group $\pm i$, $\pm j$ and $\pm k$ constituted conjugacy classes of their own and due to the commutativity of i with the other elements this result can be copied here. The character table that may be derived is Table 2.4.

Table 2.4: Character table of the biquaternion group

$\chi(K_l)$	K_1 (1)	K_2 (-1)	K_3 (i)	K_4 $(-i)$	K_5 $(\pm i_1)$	K_6 $(\pm i_2)$	K_7 $(\pm i_3)$	K_8 $(\pm ii_1)$	K_9 $(\pm ii_2)$	K_{10} $(\pm ii_3)$
$D^{(1)}$	1	1	1	1	1	1	1	1	1	1
$D^{(2)}$	1	1	1	1	1	-1	-1	1	-1	-1
$D^{(3)}$	1	1	1	1	-1	1	-1	-1	1	-1
$D^{(4)}$	1	1	1	1	-1	-1	1	-1	-1	1
$D^{(5)}$	1	1	-1	-1	1	1	1	-1	-1	-1
$D^{(6)}$	1	1	-1	-1	1	-1	-1	-1	1	1
$D^{(7)}$	1	1	-1	-1	-1	1	-1	1	-1	1
$D^{(8)}$	1	1	-1	-1	-1	-1	1	1	1	-1
$D^{(9)}$	2	-2	$2i$	$-2i$	0	0	0	0	0	0
$D^{(10)}$	2	-2	$-2i$	$2i$	0	0	0	0	0	0

Complex representation of biquaternions

In the case of biquaternions there exist eight irreducible one-dimensional representations and two irreducible two-dimensional representations. We will investigate a 2×2 complex representation of biquaternions. Choosing the first of the two two-dimensional representations and writing D instead of $D^{(9)}$ we have

$$1 \longleftrightarrow D(1) = 1_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{2.3.9}$$

$$i \longleftrightarrow D(i) = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} \tag{2.3.10}$$

For the quaternion units we apply our previous identification (2.2.17) with the Pauli matrices,

$$i_1 \longleftrightarrow i\sigma^3 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, i_2 \longleftrightarrow i\sigma^2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, i_3 \longleftrightarrow i\sigma^1 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \tag{2.3.11}$$

The representations of other elements like ii_1 follow using the relation (2.3.10), but it does turn out we do not need them, for a general biquaternion q can be represented by an irreducible 2×2 complex matrix as

$$q = c^0 + c^1 i_1 + c^2 i_2 + c^3 i_3 \longleftrightarrow D(q) = \begin{bmatrix} c^0 + c^1 i & c^2 + c^3 i \\ -c^2 + c^3 i & c^0 - c^1 i \end{bmatrix} \tag{2.3.12}$$

It is of importance to note that for biquaternions taking the quaternion conjugate no longer amounts to taking the Hermitian conjugate of its matrix representation, i.e. $D(q)^\dagger \neq D(q^\dagger)$, as was the case for quaternions in (2.2.19).

An other complex representation of biquaternions

Again, it is possible to create a reducible 4×4 representation of biquaternions out of the above irreducible 2×2 complex representation by making the following replacements in (2.3.9) and (2.3.11).

$$1 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, 0 \to \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, i \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{2.3.13}$$

Due to the complex coefficients of a biquaternion this representation is not real as the 4×4 representation (2.2.21) of a normal quaternion. It does, however, turn out to be so that we do not need to take into account the fact that the coefficients are complex and hence, we are left with

$$q = c^0 + c^1 i_1 + c^2 i_2 + c^3 i_3 \longleftrightarrow D(q) = \begin{bmatrix} c^0 & c^1 & c^2 & c^3 \\ -c^1 & c^0 & -c^3 & c^2 \\ -c^2 & c^3 & c^0 & -c^1 \\ -c^3 & -c^2 & c^1 & c^0 \end{bmatrix} \tag{2.3.14}$$

In this representation it is obvious that the vector part of – in this case – a complex quaternion is represented by an antisymmetric matrix. Below, we will see that biquaternions with only such a vector part form a special type of complex quaternions that are of more help in a formulation of electrodynamics than were the vector parts of normal quaternions.

2.3.3 Types of biquaternions

The different modes of conjugation described at the end of section 2.3.1 allow us to distinghuish several types of biquaternions. A biquaternion for which $q^* = q$ is a normal (real) quaternion, for example. For $q^\dagger = q$ it is a complex number and $q^\dagger = q^* = q$ gives us a real number. Other more interesting cases are provided by different conditions.

Hermitian biquaternions

Consider the case of the combination $q^\dagger = q^*$, for instance. Such a biquaternion has a real scalar part and the coefficients of the vector part are pure imaginary. It will be referred to as a *Hermitian biquaternion.*

Let's explain the origin of this name. The biquaternions that satisfy the property $q^\dagger = q^*$ can be written

$$q = a^0 + a^1 ii_1 + a^2 ii_2 + a^3 ii_3 = a^0 + i(a^1 i_1 + a^2 i_2 + a^3 i_3) = a^0 + i\mathbf{a} \tag{2.3.15}$$

where the coefficients a^0 to a^3 are all real and due to (2.3.12) q can be represented by a Hermitian matrix.

$$\begin{aligned} q = a^0 + a^1 ii_1 + a^2 ii_2 + a^3 ii_3 &\longleftrightarrow D(q) = \begin{bmatrix} a^0 + (a^1 i)i & a^2 i + (a^3 i)i \\ -a^2 i + (a^3 i)i & a^0 - (a^1 i)i \end{bmatrix} \\ &= \begin{bmatrix} a^0 + (a^1 i)i & -a^2 i + (a^3 i)i \\ a^2 i + (a^3 i)i & a^0 - (a^1 i)i \end{bmatrix}^* = \begin{bmatrix} a^0 + (a^1 i)i & a^2 i + (a^3 i)i \\ -a^2 i + (a^3 i)i & a^0 - (a^1 i)i \end{bmatrix}^\dagger = D(q)^\dagger \end{aligned} \tag{2.3.16}$$

For Hermitian complex quaternion we can use a straight-forward definition of a norm.

$$\begin{aligned} |q|^2 :=\ & q^\dagger q = qq^\dagger = q^* q = qq^* \\ =\ & (a^0 + a^1 ii_1 + a^2 ii_2 + a^3 ii_3)(a^0 - a^1 ii_1 - a^2 ii_2 - a^3 ii_3) := (a^0 + i\mathbf{a})(a^0 - i\mathbf{a}) \\ =\ & (a^0)^2 - \mathbf{a}\cdot\mathbf{a} \end{aligned} \tag{2.3.17}$$

where it is to be noted that $\mathbf{a}$ again denotes a vector with real coefficients.

Antisymmetric biquaternions

A different condition, namely $q^\dagger = -q$, results in a complex quaternion with only a vector part. Such a biquaternion has six independent components and we will write it as follows, using the letter A to refer to them.

$$A = \mathbf{c} = c^l i_l = (a^l + ib^l)i_l = a^l i_l + ib^l i_l = \mathbf{a} + i\mathbf{b} \tag{2.3.18}$$

As we remarked at the end of section 2.3.2, antisymmetric biquaternions may be represented by a 4×4 antisymmetric matrix, hence the name *antisymmetric biquaternion*, and in the next section we will relate them to antisymmetric matrices as a preliminary result for chapter 5.

2.3.4 Identification of four-vectors and tensors with biquaternions

Four-vectors

The Hermitian biquaternions are of help in the problem with the magnitude encountered in section 2.2.3 and therefore, we will identify four-vectors with them.

$$x = x^0 + x^1 ii_1 + x^2 ii_2 + x^3 ii_3 \longleftrightarrow x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{2.3.19}$$

This provides us with the correct squared modulus we found in (2.3.17),

$$|x|^2 := x^\dagger x = x^* x = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x^0)^2 - \mathbf{x} \cdot \mathbf{x} \tag{2.3.20}$$

In the case of Hermitian complex quaternions both modes of conjugation, i.e. quaternion and complex conjugation, may be related to the change from a contravariant to a covariant vector, because they have the same effect. Notwithstanding, the following calculation from relativistic quantum mechanics tells us which to choose. To this purpose, let us define the *momentum biquaternion*

$$p := p^0 + p^1 ii_1 + p^2 ii_2 + p^3 ii_3 = p^0 + i(p^1 i_1 + p^2 i_2 + p^3 i_3) = p^0 + i\mathbf{p} \tag{2.3.21}$$

and the *biquaternion covariant partial derivative*

$$\partial := \partial_0 - i(\partial_1 i_1 - \partial_2 i_2 - \partial_3 i_3) := \partial_0 - i\nabla \tag{2.3.22}$$

Our exercise now consists of comparing the squared modulus of the momentum in both formalisms. In the four-vector formalism

$$p^\mu p_\mu = \eta_{\mu\nu} p^\mu p^\nu = (p^0)^2 - \mathbf{p} \cdot \mathbf{p} \tag{2.3.23}$$

Filling in the usual operator definitions from quantum mechanics, $E = p^0 c \to i\hbar\frac{\partial}{\partial t}$ and $p_l \to -i\hbar\partial_l$ (or, equivalently, for the contravariant momentum four-vector, $p^l \to i\partial_l$) this gives

$$p^\mu p_\mu = -\hbar^2 \partial_0^2 + \hbar^2 \nabla \cdot \nabla \tag{2.3.24}$$

Now, making the same transition from the (contravariant) momentum biquaternion (2.3.21) to the partial derivative (2.3.22),

$$p = p^0 + i\mathbf{p} \to i\hbar\partial = i\hbar(\partial_0 - i\nabla) = i\hbar\partial_0 + \hbar\nabla \tag{2.3.25}$$

we get

$$p^\dagger p = -\hbar^2 \partial_0^2 + \hbar^2 \nabla \cdot \nabla \tag{2.3.26}$$

which is the desired result as it is the same as (2.3.24). A simple calculation shows that $p^* p$ does not provide us with this answer and therefore, we will make the following identification between covariant four-vectors and biquaternions.

$$x^\dagger = x^0 - x^1 ii_1 - x^2 ii_2 - x^3 ii_3 \longleftrightarrow x_\mu = \begin{bmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{bmatrix} \tag{2.3.27}$$

Tensors

Let us look at antisymmetric second-rank tensors again. These matrices contain six independent components and therefore, our previously defined antisymmetric biquaternions probably could be nicely identified with them. For an antisymmetric tensor A

$$A^{\nu\mu} = -A^{\mu\nu} \text{ and } A_{\nu\mu} = -A_{\mu\nu} \tag{2.3.28}$$

Something similar occurs for an antisymmetric biquaternion.

$$A^\dagger = \mathbf{c}^\dagger = (c^l i_l)^\dagger = -c^l i_l = -c^1 i_1 - c^2 i_2 - c^3 i_3 = -A \tag{2.3.29}$$

and this leads us to relate the quaternion conjugation operation with the swapping of indices of the tensors mentioned.

Using (2.3.18) we can see what the effect of complex conjugation on an antisymmetric biquaternion is.

$$A^* = (\mathbf{a} + i\mathbf{b})^* = \mathbf{a} - i\mathbf{b} \tag{2.3.30}$$

So, exactly half of the six independent components gets a minus sign. This is precisely what occurs in lowering both indices of an antisymmetric tensor as we may find by inspecting what happens to an arbitrary antisymmetric matrix upon application of the Minkowsky metric in the following way.

$$A_{\mu\nu} = A^\rho{}_\nu \eta_{\rho\mu} = \eta_{\mu\rho} A^\rho{}_\nu = \eta_{\mu\rho} A^{\rho\sigma} \eta_{\sigma\nu} = \eta_{\mu\rho}\eta_{\nu\sigma} A^{\rho\sigma} \tag{2.3.31}$$

Therefore, we can identify this operation of lowering both indices with (2.3.30), complex conjugation of an antisymmetric biquaternion.

Lowering just one of the indices seems to have no analogy in our complex quaternion formalism, because where we can treat the doubly occurring numbers in the tensors differently, in the biquaternions we cannot as they occur only once.

Despite this fact, antisymmetric biquaternions and especially their 4×4 representation (2.3.14) will prove useful in finding a complex quaternion representation of the electromagnetic field strength tensor in chapter 5 and hence, of electrodynamics.

2.4 Conclusion

In this chapter, we have investigated the properties of quaternions and biquaternions. We have explored the options to relate quaternions and complex-quaternions to four-vectors and antisymmetric tensors and from this followed a relation with Hermitian and antisymmetric biquaternions, respectively.

Chapter 3

The Dirac equation

3.1 Introduction

Now, that the formalism of quaternions and complex quaternions has been firmly established as our mathematical basis, let us look at the physics of some relativistic wave equations to which we would like to apply our gained mathematical results. In the next section we will take a look at the Klein-Gordon equation being second-order in the derivatives, which led Dirac to formulate his first-order wave equation of which we will speak in section 3.3. Discussion of the Klein-Gordon and Dirac equations will lead us naturally to the Pauli and Dirac algebras. We will treat these and their relation with the (complex) quaternions in sections 3.4 and 3.5, respectively. Conclusions will again be drawn in the final section of this chapter.

3.2 The Klein-Gordon equation

The central equation of standard quantum mechanics is the Schrödinger equation. This equation contains a first-order time derivative, while it is second order in the spatial derivatives. Due to this unequal placement of space and time it is not invariant under Lorentz transformations, i.e. for particle velocities of the order of the speed of light it gives incorrect results. To find a correct relativistic wave equation, the energy momentum relation of special relativity,

$$E^2 = \mathbf{p} \cdot \mathbf{p} c^2 + m^2 c^4 \tag{3.2.1}$$

was modified by making the replacements $E \to i\hbar\frac{\partial}{\partial t}$ and $p \to -i\hbar\nabla$. This resulted in the *Klein-Gordon equation*,

$$(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla \cdot \nabla + \frac{m^2c^2}{\hbar^2})\psi(x,t) = (\eta^{\mu\nu}\partial_\mu\partial_\nu + \frac{m^2c^2}{\hbar^2})\psi(x,t) = 0 \tag{3.2.2}$$

The Klein-Gordon equation is a relativistically invariant wave equation that describes the movement of scalar particles, that is particles with spin 0.

Using the biquaternion notation from (2.3.22) we may concisely write

$$(\partial^\dagger\partial + \frac{m^2c^2}{\hbar^2})\psi(x,t) = 0 \tag{3.2.3}$$

3.3 The Dirac equation

In relativistic quantum mechanics the *Dirac equation* is usually stated as

$$(i\gamma^\mu\partial_\mu - \frac{mc}{\hbar})\psi(x,t) = 0 \tag{3.3.1}$$

This equation gives a description of spin-$\frac{1}{2}$ particles like the electron. In the equation the constant m is the rest mass of the electron, c signifies the speed of light and $\hbar$ is the reduced Planck constant. $\psi(x,t)$

is a complex wavefunction called a *spinor* having four components. Therefore, it is implicitly assumed that a 4×4 unit matrix, 1_4, is contained in the term with all the constants.

Historically, the formulation of this equation by Paul Dirac in 1928 was of great significance, because it requires and, therefore, predicted the existence of antiparticles. Although the spinor ψ contains eight components as it consists of four complex numbers, the Dirac equation has four independent solutions due to the fact that it is a system of four coupled equations. These solutions also satisfy the Klein-Gordon equation by construction, because Dirac tried to find a square root, in some sense, of the Klein-Gordon equation. When the Dirac equation (3.3.1) holds, also

$$(i\gamma^\mu \partial_\mu + \frac{mc}{\hbar})(i\gamma^\nu \partial_\nu - \frac{mc}{\hbar})\psi(x,t) = (-\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu - \frac{m^2c^2}{\hbar^2})\psi(x,t) = 0 \tag{3.3.2}$$

To have the four components of the spinor ψ satisfy the Klein-Gordon equation (3.2.2), we require

$$\{\gamma^\mu, \gamma^\nu\} := \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \quad \text{for } \mu, \nu = 0, 1, 2, 3 \tag{3.3.3}$$

This is the so-called Dirac algebra that requires four anticommuting matrices. We will treat the Dirac algebra in more detail in section 3.5.

3.4 The Pauli algebra

Before continuing to a more detailed description of the Dirac algebra, let us note that the fact that anticommuting matrices are required by (3.3.3) is reminiscent of the Pauli algebra.

The Pauli algebra is summarized by the anticommutation relation

$$\{\sigma^l, \sigma^m\} := \sigma^l\sigma^m + \sigma^m\sigma^l = 2\delta^{lm}1_2 \quad \text{for } l, m = 1, 2, 3 \tag{3.4.1}$$

and for completeness we state the commutation relation here, as well

$$[\sigma^l, \sigma^m] := \sigma^l\sigma^m - \sigma^m\sigma^l = 2i\epsilon^{lmn}\sigma^n \quad \text{for } l, m = 1, 2, 3 \tag{3.4.2}$$

Note that the above relations are equivalent to the earlier stated (2.2.16) and they imply the Pauli algebra requires three anticommuting quantities.

Conventionally, the Pauli matrices, having to do with spin in quantum mechanics, are taken to be

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.4.3}$$

as stated before in (2.2.15) and indeed these 2×2 complex matrices satisfy the relations (3.4.1) and (3.4.2) given above and other properties that follow from it, like

$$(\sigma^1)^2 = (\sigma^2)^2 = (\sigma^3)^2 = -i\sigma^1\sigma^2\sigma^3 = 1_2 \tag{3.4.4}$$

Now, let us look at possibilities to use (bi)quaternions as elements to satisfy the Pauli algebra.

3.4.1 Identification with quaternions

In equation (2.2.17) we saw it is possible to relate the Pauli matrices and the quaternion units via the identification

$$i \longleftrightarrow i\sigma^3, j \longleftrightarrow i\sigma^2, k \longleftrightarrow i\sigma^1 \tag{3.4.5}$$

One can check that these quantities of which i, j and k anticommute indeed satisfy the Pauli algebra (3.4.1) – and this makes sense, because (3.4.5) provides us with an isomorphism between the quaternion group and the Pauli matrices. Working out an example related to (3.4.4),

$$-i\sigma^1\sigma^2\sigma^3 = -i(i\sigma^1)(i\sigma^2)(i\sigma^3)i \longleftrightarrow kji = -i^2 = 1 \longleftrightarrow 1_2 \tag{3.4.6}$$

Other identifications are possible, however. First of all, similarity transformations allow for different quantities being identified with the Pauli matrices. For instance, a certain rotation of the quaternion units represented by $-j(\ )j$ results in

$$-i \longleftrightarrow i\sigma^3, j \longleftrightarrow i\sigma^2, -k \longleftrightarrow i\sigma^1 \tag{3.4.7}$$

that perfectly fits the algebra defined above.

Secondly, the anticommutativity of quaternion multiplication allows for pretty strange identifications involving right multiplication. Here, $a(\ )b$ means left multiplication by a and right multiplication by b. Note the reversed order of the Pauli matrices.

$$-i(\ )k \longleftrightarrow \sigma^1, i(\ )j \longleftrightarrow \sigma^2, -i(\ )i \longleftrightarrow \sigma^3 \tag{3.4.8}$$

One should check the relations that are to be satisfied, of course. An example, where ψ is inserted to distinguish left and right multiplication is given here.

$$\sigma^1\sigma^2\psi \longleftrightarrow \sigma^1(i\psi j) \longleftrightarrow -ii\psi jk = i(-i)\psi i \longleftrightarrow i\sigma^3\psi \tag{3.4.9}$$

We will encounter this representation in the next chapter.

3.4.2 Identification with biquaternions

If we are able to directly relate the Pauli matrices with the quaternion units, we must be able to relate them with elements from the biquaternion group as well. The resemblance of the multiplicative properties of the biquaternions in (2.3.3) with the relation of the Pauli matrices (2.2.16) supports this thought. In fact, the identification of some of the biquaternion units with the Pauli matrices can easily be summarized in our biquaternion notation.

$$i_l \longleftrightarrow i\sigma^{4-l} \qquad \text{for } l = 1, 2, 3 \tag{3.4.10}$$

This does not seem to add anything to our spectrum of representations, because we have not used the fact that the biquaternions form a larger group than the quaternions do. The biquaternions do offer us possibilities the quaternions do not. Multiplication of the above identification (3.4.10) with $-i$, for instance, gives

$$-ii_l \longleftrightarrow \sigma^{4-l} \qquad \text{for } l = 1, 2, 3 \tag{3.4.11}$$

and as one can check the quantities $-ii_3$, $-ii_2$ and $-ii_1$ indeed satisfy the Pauli algebra.

Here, similarity transformations and identifications using right multiplication are possible as well, however we will see that the biquaternion elements also satisfy the Dirac algebra that is more general than the Pauli algebra. Therefore, we will treat them in the following section.

3.5 The Dirac algebra

Let us return to the Dirac algebra that we encountered in section 3.3,

$$\{\gamma^\mu, \gamma^\nu\} := \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \qquad \text{for } \mu, \nu = 0, 1, 2, 3 \tag{3.5.1}$$

Here the matrix η is the Minkowski metric again, the diagonal matrix with signature $(+, -, -, -)$.

This anticommutation relation requires four anticommuting elements, whereas the Pauli matrices provide us with only three. Therefore, we need to take recourse to square matrices with dimensions higher than that of the Pauli matrices. Usually, they are 4×4 matrices that are taken to contain the 2×2 Pauli matrices. In the *standard* or *Dirac representation* the matrices γ^μ are given by

$$\gamma^0 = \begin{bmatrix} 1_2 & \\ & -1_2 \end{bmatrix}, \gamma^l = \begin{bmatrix} & \sigma^l \\ -\sigma^l & \end{bmatrix} \qquad \text{for } l = 1, 2, 3 \tag{3.5.2}$$

In the *chiral* or *Weyl basis* they can be written

$$\gamma^0 = \begin{bmatrix} & 1_2 \\ 1_2 & \end{bmatrix}, \gamma^l = \begin{bmatrix} & \sigma^l \\ -\sigma^l & \end{bmatrix} \qquad \text{for } l = 1, 2, 3 \tag{3.5.3}$$

A representation that is called the *alternative choice* has a particularly elegant formulation in terms of quaternions as we will see below.

$$\gamma^0 = \begin{bmatrix} & 1_2 \\ 1_2 & \end{bmatrix}, \gamma^l = \begin{bmatrix} i\sigma^l & \\ & -i\sigma^l \end{bmatrix} \quad \text{for } l = 1, 2, 3 \tag{3.5.4}$$

In later chapters we will use the first two of these representations and encounter its quaternionic counterpart, which we will describe below, as well.

Using the previous section we can write down several quaternion and biquaternion representation of the gamma matrices and we will state some below.

3.5.1 Quaternions

From the previous it is clear the Pauli matrices are used in representations of the gamma matrices. This allows us to use section 3.4 to replace the Pauli matrix blocks in the gamma matrices by quaternions or biquaternions. Thereby we make a transition from the usual complex 4×4 matrices to quaternion 2×2 matrices. Due to this, it is to be expected that we can halve the number of components of the Dirac equation by using a quaternion formalism.

The mentioned alternative representation (3.5.4) of the γ^μ or *Dirac matrices* can be translated to a 2×2 quaternion representation using (3.4.5).

$$\gamma^0 = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \gamma^1 = \begin{bmatrix} k & \\ & -k \end{bmatrix}, \gamma^2 = \begin{bmatrix} j & \\ & -j \end{bmatrix}, \gamma^3 = \begin{bmatrix} i & \\ & -i \end{bmatrix} \tag{3.5.5}$$

Of course, different representations may be used. The identification (3.4.8) gives us the somewhat strange, but working

$$\gamma^0 = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \gamma^1 = \begin{bmatrix} & -i(\ )k \\ i(\ )k & \end{bmatrix}, \gamma^2 = \begin{bmatrix} & i(\ )j \\ i(\ )j & \end{bmatrix}, \gamma^3 = \begin{bmatrix} & -i(\ )i \\ i(\ )i & \end{bmatrix} \tag{3.5.6}$$

Remember, that $i(\ )j$ means multiplication by i from the left and by j from the right as was illustrated in (3.4.9).

3.5.2 Biquaternions

As we have seen in section 3.4 it was perfectly possible to relate biquaternions with the Pauli matrices. Hence, we may substitute the relations (3.4.10) in the mentioned gamma matrix representation (3.5.4). Also, (3.4.11) may be substituted in both the Dirac representation (3.5.2) and the Weyl representation (3.5.3).

Nevertheless, we will not pursue that approach and make use of the fact that the biquaternion group as a larger group offers us more possibilities. Taking, for instance, the operators

$$i(\ )i_1, ii_1(\ )i_2, ii_2(\ )i_2, ii_3(\ )i_2 \tag{3.5.7}$$

we see that we have four anticommuting elements of which the first leaves us with unity when squared and the other three provide us with -1 upon squaring. This means these four elements are equivalent to any Dirac matrix representation we may find, i.e. we have found a one-dimensional biquaternion representation of the gamma matrices.

Obviously, replacement of right multiplication by i_2 in (3.5.7) by i_3 leaves us with another representation of the γ^μ matrices and of course, even that does not exhaust our options. A more detailed account of the options will be given when we speak of Conway's biquaternion formulation of the Dirac equation in section 4.5.

3.6 Conclusion

After introducing the Klein-Gordon and Dirac equations, we were naturally led to the Dirac algebra as a restriction of the γ^μ matrices. We found 2×2 quaternion and 1×1 biquaternion representations of these matrices. Also, quaternion and complex quaternion representations of the Pauli matrices were found.

Chapter 4

Quaternion formulation of the Dirac equation

4.1 Introduction

In the previous chapters we aimed to build up our knowledge concerning quaternions and complex quaternions and we have looked at the Dirac equation in particular. In this chapter we will try to connect these different topics by applying the quaternion formalism to the Dirac equation. In the next section of this chapter we will simply do this by making a translation of the Dirac equation as usually stated to a quaternion formalism. The third section will be about a different approach to the Dirac equation, in which a two-component biquaternion wave equation will be treated. Sections 4.4 and 4.5 will treat a one-component biquaternion equation. The last section of this chapter contains some concluding remarks.

4.2 Translation to quaternions

In translating the Dirac equation (3.3.1) to the quaternion formalism we will choose the Dirac and Weyl representations mentioned in section 3.5 of the Dirac algebra to write out the different components of the equation. Than, reading the complex unit in the Dirac as the quaternion unit i, we will use a "symplectic" way of writing our four-component complex spinor as a two-component quaternion spinor.

$$\psi(x,t) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} \longleftrightarrow \begin{bmatrix} \psi_1 + \psi_2 j \\ \psi_3 + \psi_4 j \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \phi(x,t) \tag{4.2.1}$$

The Dirac representation

Now, starting with the standard representation of the Dirac equation (3.5.2), applying our identification of the four-component with the two-component spinor (4.2.1) and collecting as many terms as possible, we get

$$
\begin{aligned}
(i\gamma^{\mu}\partial_{\mu} - \frac{mc}{\hbar})\psi(x,t) &= (i\gamma^0\partial_0 + i\gamma^1\partial_1 + i\gamma^2\partial_2 + i\gamma^3\partial_3 - \frac{mc}{\hbar}1_4)\psi(x,t) \\
&= \begin{bmatrix} i\partial_0 - \frac{mc}{\hbar} & 0 & i\partial_3 & i\partial_1 + \partial_2 \\ 0 & i\partial_0 - \frac{mc}{\hbar} & i\partial_1 - \partial_2 & -i\partial_3 \\ -i\partial_3 & -i\partial_1 - \partial_2 & -i\partial_0 - \frac{mc}{\hbar} & 0 \\ -i\partial_1 + \partial_2 & i\partial_3 & 0 & -i\partial_0 - \frac{mc}{\hbar} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} \\
&\longleftrightarrow \begin{bmatrix} (i\partial_0 - \frac{mc}{\hbar})(\psi_1 + \psi_2 j) \\ (-i\partial_0 - \frac{mc}{\hbar})(\psi_3 + \psi_4 j) \end{bmatrix} + \\
&\quad \begin{bmatrix} i\partial_3\psi_3 + (i\partial_1 - \partial_2)\psi_3 j + (i\partial_1 + \partial_2)\psi_4 - i\partial_3\psi_4 j \\ -i\partial_3\psi_1 - (i\partial_1 - \partial_2)\psi_1 j - (i\partial_1 + \partial_2)\psi_2 + i\partial_3\psi_2 j \end{bmatrix}
\end{aligned} \tag{4.2.2}
$$

Here we observe a problem. A quaternion equivalent of the γ^0 matrix can easily be extracted from our work as

$$
\begin{bmatrix} (i\partial_0 - \frac{mc}{\hbar})(\psi_1 + \psi_2 j) \\ (-i\partial_0 - \frac{mc}{\hbar})(\psi_3 + \psi_4 j) \end{bmatrix} = i\partial_0 \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} - \frac{mc}{\hbar} \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} \tag{4.2.3}
$$

but due to the right multiplication of j there seems to be no way to do the same for the other gamma matrices. It is, however, exactly right multiplication that helps us solve this problem and allows us to collect more terms. Observe, for instance, that

$$
\begin{aligned}
i\partial_1\psi_3 j &= \partial_1\psi_3 k & i\partial_1\psi_4 &= \partial_1\psi_4 jk \\
-\partial_2\psi_3 j &= -\partial_2\psi_3 j & \partial_2\psi_4 &= -\partial_2\psi_4 jj \\
i\partial_3\psi_3 &= \partial_3\psi_3 i & -i\partial_3\psi_4 j &= \partial_3\psi_4 ji
\end{aligned} \tag{4.2.4}
$$

This allows us to rewrite the second two-row matrix in (4.2.2) as

$$
\begin{aligned}
&\begin{bmatrix} \partial_3\psi_3 i + \partial_1\psi_3 k - \partial_2\psi_3 j + \partial_1\psi_4 jk - \partial_2\psi_4 jj + \partial_3\psi_4 ji \\ -\partial_3\psi_3 i - \partial_1\psi_3 k + \partial_2\psi_3 j - \partial_1\psi_4 jk + \partial_2\psi_4 jj - \partial_3\psi_4 ji \end{bmatrix} \\
=& \begin{bmatrix} (\partial_1(\ )k - \partial_2(\ )j + \partial_3(\ )i)(\psi_3 + \psi_4 j) \\ (-\partial_1(\ )k + \partial_2(\ )j - \partial_3(\ )i)(\psi_1 + \psi_2 j) \end{bmatrix} \\
=& \begin{bmatrix} & \partial_1(\ )k - \partial_2(\ )j + \partial_3(\ )i \\ -\partial_1(\ )k + \partial_2(\ )j - \partial_3(\ )i & \end{bmatrix} \begin{bmatrix} \psi_1 + \psi_2 j \\ \psi_3 + \psi_4 j \end{bmatrix}
\end{aligned} \tag{4.2.5}
$$

Summarizing the above results we state our quaternion translation of the Dirac, equation

$$
\begin{bmatrix} i\partial_0 - \frac{mc}{\hbar} & \partial_1(\ )k - \partial_2(\ )j + \partial_3(\ )i \\ -\partial_1(\ )k + \partial_2(\ )j - \partial_3(\ )i & -i\partial_0 - \frac{mc}{\hbar} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.2.6}
$$

or

$$
(i\gamma^{\mu}\partial_{\mu} - \frac{mc}{\hbar})\phi(x,t) = 0 \tag{4.2.7}
$$

where $\phi(x,t)$ is a two-component quaternion spinor and the gamma matrices are the following 2×2 matrices with quaternion coefficients. Note that an extra factor of $-i$ has been introduced in γ^l.

$$
\gamma^0 = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \gamma^1 = i(\ )k \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \gamma^2 = -i(\ )j \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \gamma^3 = i(\ )i \begin{bmatrix} & 1 \\ -1 & \end{bmatrix} \tag{4.2.8}
$$

These entities do indeed satisfy the Dirac algebra as noted in (3.5.6).

This approach does not seem to add anything at all to our physical understanding of the Dirac equation and in this approach we have found no principles that are more general than that of the gamma matrices in the normal approach. We did find a quaternion representation of the Dirac equation, but unfortunately it is not simple due to the appearance of right multiplication. We will, however, apply the approach used above again and find a quaternion translation of the Dirac equation in the Weyl representation as it forms a natural transition to the next section.

Note that it is clear that the two quaternion components of our spinor ϕ satisfy the Klein-Gordon equation as we found a matrix representation of gamma matrices and they satisfy the Dirac algebra.

The Weyl representation

Following the approach of the previous section we get for the Weyl representation

$$\begin{bmatrix} -\frac{mc}{\hbar} & i\partial_0 + \partial_1(\ )k - \partial_2(\ )j + \partial_3(\ )i \\ i\partial_0 - \partial_1(\ )k + \partial_2(\ )j - \partial_3(\ )i & -\frac{mc}{\hbar} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.2.9}$$

As expected, only the γ^0 matrix differs from the previous case being

$$\gamma^0 = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} \tag{4.2.10}$$

In anticipation of the following we would like to define $\nabla' = i\partial_1(\ )k - i\partial_2(\ )j + i\partial_3(\ )i$ such that we can write (4.2.9) as

$$\begin{aligned} i(\partial_0 + \nabla')\phi_1 &= \frac{mc}{\hbar}\phi_2 \\ i(\partial_0 - \nabla')\phi_2 &= \frac{mc}{\hbar}\phi_1 \end{aligned} \tag{4.2.11}$$

4.3 Towards a biquaternion Dirac equation

It is our aim to arrive at a formulation of the Dirac equation that is as simple as possible, but still contains all the expected physics – and perhaps some unexpected physics – in a straight-forward way. The translation made in the previous section is still an equation having components and the physics it conveys is of course still the same as that of the usual formulation.

We would like to arrive at an elegant one-component Dirac equation and due to biquaternion representations of the γ^μ matrices like (3.5.7) at least the target of finding a single equation is expected to be reached. In the next section we will use the translation approach of section 4.2 to find such an equation. Again, such a translation will probably not contain any more than the usual physics. That is why we would like to draw attention to the two-component biquaternion formulation of the Dirac equation of James Edmonds [6] that is also worked out in more detail by Stefano de Leo [7]. Again using our biquaternion covariant derivative (2.3.22) and noting the biquaternion character of the spinor $\chi(x,t)$ we may state Edmonds' result as

$$\begin{aligned} i\partial\chi_1 &= \frac{mc}{\hbar}\chi_2 \\ i\partial^\dagger\chi_2 &= \frac{mc}{\hbar}\chi_1 \end{aligned} \tag{4.3.1}$$

His formulation looks very much like our quaternion translation of the Weyl representation (4.2.11) apart from the biquaternion character of this equation.

Here, we can obtain the Klein-Gordon equation (in biquaternion notation, (3.2.3)) with even more ease when using both previous equations.

$$\begin{aligned} i\partial\frac{mc}{\hbar}\chi_1 = i\partial i\partial^\dagger\chi_2 &= -\partial^\dagger\partial\chi_2 \\ &= (\frac{mc}{\hbar})^2\chi_2 \end{aligned} \tag{4.3.2}$$

Although, the two different biquaternion components of the wavefunction χ Edmonds uses obviously satsify the Klein-Gordon equation, as was shown for one of the components, there are some things that need to be remarked. Edmonds' version of the Dirac equation (4.3.1) cannot be equivalent to the Dirac equation because that would require it to have eight-component complex spinors when stated as usual. This leads Edmonds to say that the number of independent solutions of his Dirac equation doubles to eight and that due to this it contains "new physics". That is, it allows for more degrees of freedom than usually encountered. We question this approach not only due to De Leo, who argues that the newly found solution arise simply because of the use of a reducible representation [7]. It is also due to the article of Dirk Schuricht and Martin Greiter [8] that gives an in-depth account of all the problems encountered when investigating a two-component biquaternion Dirac equation. Besides, it seems unphysical to simply allow for the components of the wavefunction to contain more numbers as it may be possible to take an other hypercomplex number formalism and say the number of solutions quadruples.

This leads us to leave Edmonds' speculation and investigate two forms of a one-component biquaternion Dirac equation in the following sections.

4.4 Translation to biquaternions

Translating the Dirac equation to a form where biquaternions are used is of course not much different than translating it to quaternions like we did in section 4.2. We will write the usual four-component complex spinor as a one-component biquaternion spinor using the following "symplectic" way of writing it. Note that the complex numbers ψ_1 to ψ_4 now contain the complex unit i.

$$\psi(x,t) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} \longleftrightarrow \begin{bmatrix} \psi_1 + \psi_2 i_1 + \psi_3 i_2 + \psi_4 i_3 \end{bmatrix} = \begin{bmatrix} \chi_1 \end{bmatrix} = \chi(x,t) \tag{4.4.1}$$

We chose the Weyl representation to work out the usual Dirac equation, like we did for the Dirac representation in (4.2.2). Then, encountering the problem of right multiplication again and solving this by making observations similar to those in (4.2.4) we arrive at the following biquaternion Dirac equation.

$$(i(i_1\partial_0(\ )i_3 - i_3\partial_1 + ii_1\partial_2(\ )i_2 + i_2\partial_3) + \frac{mc}{\hbar})\chi(x,t) = 0 \tag{4.4.2}$$

We may check that $i_1(\ )i_3$, i_3, $ii_1(\ )i_2$ and i_2 indeed satisfy the Dirac algebra (3.3.3). Note that we can also easily check this mentally as i_1, i_2 and i_3 anticommute and their squares are -1. Now, our equation (4.4.2) does not look particularly enlightening. It is still built upon the Dirac algebra requirement to have the Dirac equation wave function components satisfy the Klein-Gordon equation. Comparing the representation found now with (3.5.7), however, shows us the biquaternions may provide a very flexible candidate for formulating the Dirac equation. Let us look into this, now.

4.5 Biquaternion operator approach

Before, we began with the usual Dirac equation with its associated Dirac algebra to find quaternion and biquaternion formulations of the Dirac equation. In this scheme the (complex) quaternions are simply used as a representation of the Dirac matrices, γ^0 to γ^3. Due to this, we may say that the found quaternion formulations of the Dirac equation are less general than the usual formulation (3.3.1). Also, it is obvious that we must be able to find all the usual results like plane wave solutions, nonrelativistic approximations, etcetera in this way. If we succeed in this, this may be seen as a triumph of the quaternion formalism, but this will only be the case if the usual results can be obtained in a more intuitive way. We argue that this is so, because it takes some time to get used to both the four-vector and quaternion formalisms and in the former the usual results are well-formulated already.

To investigate whether the quaternion formalism has some intuitiveness to offer, let us look at Arthur Conway's [4] approach instead of taking the Dirac algebra as a starting point. He was one of the first to try to apply a quaternion formalism to the Dirac equation. In his 1937 paper on this topic he suggests a method to derive a biquaternion Dirac equation. Essentially, he did this by trying the same thing as Dirac did: finding an expression that is in some sense the square root of the Klein-Gordon equation.

We will use his approach to find a one-component biquaternion equation that is equivalent to the Dirac equation and see that it is in fact equivalent to our biquaternion translation (4.4.2), as well. Conway seeks to derive an equation in which there are five anticommuting operators, namely one for the time-derivative, three for the spatial derivatives and one for the constant operator $\frac{mc}{\hbar}$.

To this end, he observes that there is a limited number of ways in which one can multiply by the quaternion units

$$\begin{array}{lll} 1(\ )i_1 & 1(\ )i_2 & 1(\ )i_3 \\ i_1(\ )1 & i_2(\ )1 & i_3(\ )1 \\ i_1(\ )i_1 & i_1(\ )i_2 & i_1(\ )i_3 \\ i_2(\ )i_1 & i_2(\ )i_2 & i_2(\ )i_3 \\ i_3(\ )i_1 & i_3(\ )i_2 & i_3(\ )i_3 \end{array} \tag{4.5.1}$$

and one can select exactly five anticommuting elements out of this. This convenient observation permits us to make the above-mentioned operators anticommute in for instance, the following way.

$$\partial_0(\ )i_1,\ i_1\partial_1(\ )i_2,\ i_2\partial_2(\ )i_2,\ i_3\partial_3(\ )i_2 \text{ and } \frac{mc}{\hbar}(\ )i_3 \tag{4.5.2}$$

Noting that we defined our biquaternion gradient as

$$\nabla = i_1\partial_1 + i_2\partial_2 + i_3\partial_3 \tag{4.5.3}$$

we will use (4.5.2) to write the biquaternion Dirac equation

$$(\partial_0(\ )i_1 + \nabla(\ )i_2 + \frac{mc}{\hbar}(\ )i_3)\chi(x,t) = \partial_0\chi i_1 + \nabla\chi i_2 + \frac{mc}{\hbar}\chi i_3 = 0 \tag{4.5.4}$$

Note that repetition of the operator used here results in the Klein-Gordon equation (3.2.3) due to anticommutativity of the different terms.

$$(\partial_0(\ )i_1 + \nabla(\ )i_2 + \frac{mc}{\hbar}(\ )i_3)^2\chi(x,t) = (-\partial_0^2 + \nabla\cdot\nabla - \frac{m^2c^2}{\hbar^2})\chi(x,t) = 0 \tag{4.5.5}$$

In Conway's approach no use is made of the Dirac algebra, but the requirement of finding anticommuting operators suggests there exists a link between them. Comparison of the elements (4.5.1) with our biquaternion translation of Dirac's equation (4.4.2) uncovers this relation. We may simply select some anticommuting elements of (4.5.1) and after multiplication by i in some cases arrive at a biquaternion representation of the Dirac matrices. This was the more detailed treatment referred to at the end of section 3.5.

Now, selecting some specific quantities from (4.5.1) of which all but one occured in our translation (4.4.2), we arrive at an other biquaternion Dirac equation. Note that here we have taken the Dirac algebra as our starting point again, as we cannot simply take the square of the operator as in Conway's approach to arrive at the Klein-Gordon equation.

$$(i(i_1\partial_0(\ )i_1 - i_3\partial_1 + ii_1\partial_2(\ )i_2 + i_2\partial_3) + \frac{mc}{\hbar})\chi(x,t) = 0 \tag{4.5.6}$$

This equation will return in chapter 6 on the physical aspects of the Dirac equation because the occurence of the same quaternion unit on both sides of ∂_0 will be required there.

4.6 Conclusion

In this chapter we have applied the formalism of (complex) quaternions to the Dirac equation and we found several suitable formulations. First of all, a two-component quaternion formulation of the Dirac equation was derived. Also, a two-component biquaternion equation was investigated, but due to some problems it was discarded. After that, we looked at one-component biquaternion formulations of Dirac's equation, which certainly had some elegance to them. We will investigate these equivalent formulations in chapter 6.

Chapter 5

Electromagnetism

5.1 Introduction

In earlier chapters we hinted at the complex quaternions providing us with a useful representation of the electromagnetic field strength tensor. As a sidestep to show the elegance of quaternions in a formulation of electrodynamics and as preperation for the next chapter, let us look at such a representation. To work this out properly, let us mention the Maxwell equation in the next section and formulate them in both the usual four-vector formalism and utilizing complex quaternions. After that, in section 5.3, let us introduce the four-potential and state some relations concerning the Lagrangian and Hamiltionian. In the final section some conclusions will be drawn.

5.2 Maxwell's equations

In the CGS system of units, Maxwell's equations may be stated as

$$\begin{aligned} &\nabla \cdot \mathbf{E} = c\rho \qquad &\nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{c\partial t} = \mathbf{j} \\ &\nabla \cdot \mathbf{B} = 0 \qquad &\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{c\partial t} = 0 \end{aligned} \tag{5.2.1}$$

5.2.1 Four-vector formalism

When one looks at electrodynamics from a special relativity point of view, an antisymmetric tensor is introduced to summarize the Maxwell equations and to be able to write them in relativistic form.

$$F^{\mu\nu} = -F^{\nu\mu} = \begin{bmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{bmatrix} \tag{5.2.2}$$

The Maxwell equation are now equivalent to the following two equation where μ, ν and ρ all may attain the values 0 to 3.

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{5.2.3}$$

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0 \tag{5.2.4}$$

5.2.2 Biquaternion formalism

Previously, in section 2.3 we related antisymmetric tensors with antisymmetric biquaternions and encouraged by Lambek we will look at the following quantity (cf. [11] and the references therein).

$$F = \mathbf{B} + i\mathbf{E} \tag{5.2.5}$$

Using our biquaternion covariant derivative $\partial = \partial_0 - i\nabla$ and applying it to the field strength biquaternion F we get

$$\begin{aligned}\partial F &= (\partial_0 - i\nabla)(\mathbf{B} + i\mathbf{E}) = \partial_0\mathbf{B} + i\partial_0\mathbf{E} - i\nabla\mathbf{B} + \nabla\mathbf{E} \\ &= \partial_0\mathbf{B} + i\partial_0\mathbf{E} + i\nabla\cdot\mathbf{B} - i\nabla\times\mathbf{B} - \nabla\cdot\mathbf{E} + \nabla\times\mathbf{E} \\ &= -\nabla\cdot\mathbf{E} + i\nabla\cdot\mathbf{B} - i(\nabla\times\mathbf{B} - \partial_0\mathbf{E}) + (\nabla\times\mathbf{E} + \partial_0\mathbf{B}) \\ &= -c\rho - i\mathbf{j}\end{aligned} \tag{5.2.6}$$

Hence, when we define the four-current as the Hermitian biquaternion

$$j = c\rho + i\mathbf{j} \tag{5.2.7}$$

we find a formulation in which the Maxwell equations are condensed to one equation

$$\partial F + j = 0 \tag{5.2.8}$$

5.3 Solutions of Maxwell's equations and other properties

In the four-vector formalism the four-potential A^μ is introduced to solve the equation (5.2.4) in the following way

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.3.1}$$

We can do something similar now, when we introduce the four-potential

$$A = c\phi + i\mathbf{A} \tag{5.3.2}$$

and apply the usual relations of the four-potential with the electric and magnetic fiels

$$\mathbf{B} = \nabla\times\mathbf{A}, \mathbf{E} = -\nabla c\phi - \partial_0\mathbf{A}$$

If we remember the effect of quaternion conjugation from (2.3.29) we may arrive at result that is analogous to (5.3.1).

$$F = \frac{1}{2}((\partial A) - (\partial A)^\dagger) \tag{5.3.3}$$

Note that the four-potential as was introduced (5.3.2) will return in the next chapter.

With a biquaternion formalism we may arrive at more expressions that are akin to the usual ones. We may for instance express the Lagrangian in terms of our antisymmetric biquaternion F. As a preliminary result, let us compute some quantities related to F.

$$F^2 = (\mathbf{B} + i\mathbf{E})^2 = -\mathbf{B}\cdot\mathbf{B} + \mathbf{E}\cdot\mathbf{E} - 2i\mathbf{E}\cdot\mathbf{B} \tag{5.3.4}$$

$$(F^*)^2 = (\mathbf{B} - i\mathbf{E})^2 = -\mathbf{B}\cdot\mathbf{B} + \mathbf{E}\cdot\mathbf{E} + 2i\mathbf{E}\cdot\mathbf{B} \tag{5.3.5}$$

and

$$F^{\dagger *}F = -F^*F = -(\mathbf{B} - i\mathbf{E})(\mathbf{B} + i\mathbf{E}) = \mathbf{B}\cdot\mathbf{B} + \mathbf{E}\cdot\mathbf{E} - 2i\mathbf{B}\times\mathbf{E} \tag{5.3.6}$$

$$FF^{\dagger *} = -FF* = -(\mathbf{B} + i\mathbf{E})(\mathbf{B} - i\mathbf{E}) = \mathbf{B}\cdot\mathbf{B} + \mathbf{E}\cdot\mathbf{E} + 2i\mathbf{B}\times\mathbf{E} \tag{5.3.7}$$

Now, as proposed by Schuricht and Greiter [10], we may write our Lagrangian for free space as

$$\mathcal{L} = \frac{1}{4}(F^2 + (F^*)^2) \tag{5.3.8}$$

It may be clear by now that the biquaternion description of the Maxwell equations has a certain elegance to it. Probably, it could be extended to cover more similar results of which the Hamiltonian might be one. For the Hamiltonian we could propose the following definition after looking carefully at equations (5.3.4) to (5.3.7).

$$\mathcal{H} = \frac{1}{4}(F^{\dagger *}F + FF^{\dagger *}) \tag{5.3.9}$$

5.4 Conclusion

In the current chapter, an elegant formulation of electrodynamics was found in which Maxwell's equations were reduced to a single biquaternion equation. Also, the Lagrangian has been formulated in terms of the antisymmetric field strength biquaternion that was introduced in this chapter and a formulation for the Hamiltonian has been proposed. The four-potential that has been defined will return in the next chapter.

Chapter 6

Physical aspects of the biquaternion Dirac equation

6.1 Introduction

Now that we have formulated a biquaternion wave equation, let us show that it is equivalent to the Dirac equation by looking at the physics behind the equation. Up to this point we have spoken about reformulating Dirac's equation in terms of quaternions and complex quaternions. In section 4.5 we formulated equivalent biquaternion wave equations in (4.5.6) and (4.5.4). In the following section we will use the former of these equations to perform the so-called nonrelativistic approximation of the Dirac equation. In section 6.3, where we will speak of plane wave solutions, the latter will prove useful. Again, in the last section of this chapter some conclusions will be presented.

6.2 Coupling to an electromagnetic field

Dirac's equation (3.3.1) describes free electrons, i.e. electrons that experience no potential from, for instance, an electric or magnetic field. In quantum mechanics, the transition from the energy and momentum observables to the following operators is made to extract energy and momentum eigenvalues from a wave function.

$$\begin{aligned} E &\to i\hbar c\partial_0 \\ p_l &\to -i\hbar\partial_l \end{aligned} \tag{6.2.1}$$

When electromagnetic fields are considered the scheme of *minimal coupling* is applied, usually. We do so by adding some simple terms to (6.2.1). Here A is the four-potential introduced in (5.3.2) in the previous chapter.

$$\begin{aligned} E &\to i\hbar c\partial_0 - eA_0(x) = i\hbar c(\partial_0 + \frac{ie}{\hbar c}A_0(x)) \\ p_l &\to -i\hbar\partial_l + \frac{e}{c}A_k(x) = -i\hbar(\partial_l + \frac{ie}{\hbar c}A_k(x)) \end{aligned} \tag{6.2.2}$$

so that in the Dirac equation we may make the following replacement.

$$i\partial_\mu \to i(\partial_\mu + \frac{ie}{\hbar c}A_\mu(x)) \tag{6.2.3}$$

Writing the previous in biquaternion notation, we have

$$i\partial \to i(\partial + \frac{ie}{\hbar c}A^\dagger(x)) \tag{6.2.4}$$

6.2.1 The nonrelativistic approximation

When making the replacement (6.2.4) in our biquaternion equation (4.5.6) we get

$$(i(i_1(\partial_0+\frac{ie}{\hbar c}A_0)(\ )i_1-i_3(\partial_1-\frac{ie}{\hbar c}A^1)+ii_1(\partial_2-\frac{ie}{\hbar c}A^2)(\ )i_2+i_2(\partial_3-\frac{ie}{\hbar c}A^3))+\frac{mc}{\hbar})\chi(x,t)=0 \quad (6.2.5)$$

Now, as we see, this is a one-component equation. In making the nonrelativistic approximation, we wish to distinguish between the small and large components of the spinor, however. In the usual approach, the four-component complex spinor is written as two complex two-component spinors for this purpose. Here we will take recourse to the following trick: we will write the biquaternion spinor χ as the sum of two spinors ϕ_1 and ϕ_2. The connection with the usual spinor components ψ_1 to ψ_4 is also given.

$$\chi=\phi_1+\phi_2 i_2=(\psi_1+\psi_2 i_1)+(\psi_3+\psi_4 i_1)i_2 \quad (6.2.6)$$

We will make the assumption that our wavefunction χ is stationary so that we may write

$$\chi(x,t)=e^{-\frac{i}{\hbar}Et}\chi(x) \quad (6.2.7)$$

Then, following the usual approach, we arrive at the following two equations.

$$(E-eA_0-mc^2)\phi_1=\hbar c\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_2 i_2 \quad (6.2.8)$$

$$(E-eA_0+mc^2)\phi_2=-i\hbar c\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_1 i_2 \quad (6.2.9)$$

Here, we have used the following definitions for the gamma matrices γ^1, γ^2 and γ^3.

$$\gamma^1=-ii_3,\gamma^2=-i_1(\ )i_2,\gamma^3=ii_2 \quad (6.2.10)$$

Now, we will make an approximation in neglecting the kinetic and potential energy with respect to the rest mass of the electron. To this purpose, we define $\bar{E}=E-mc^2$ and note that this would not have been possible in an unambiguous way if we had chosen the representations (4.4.2) or (4.5.4). This allows us to rewrite (6.2.8) and (6.2.9) to

$$(\bar{E}-eA_0)\phi_1=\hbar c\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_2 i_2 \quad (6.2.11)$$

$$(\bar{E}-eA_0+2mc^2)\phi_2=-\hbar c\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_1 i_2 \quad (6.2.12)$$

Making our approximation, $\bar{E}-eA_0<<2mc^2$, we are able to find the component ϕ_2 in terms of ϕ_1.

$$\phi_2=-i\frac{\hbar c}{2mc^2}\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_1 i_2 \quad (6.2.13)$$

This finally results in an equation for ϕ_1, namely

$$(\bar{E}-eA_0)\phi_1=\frac{\hbar^2}{2m}[\gamma\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})]^2\phi_1 \quad (6.2.14)$$

Explicitly working out the expression in square brackets using the definition of the gamma matrices (6.2.10) and the relation $\mathbf{B}=\nabla\times\mathbf{A}$ provides us with the following result.

$$\bar{E}\phi_1=eA_0\phi_1-\frac{\hbar^2}{2m}(\nabla-\frac{ie}{\hbar c}\mathbf{A})\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})\phi_1-\frac{e\hbar}{2mc}(i\gamma^2\gamma^3B^1+i\gamma^3\gamma^1B^2+i\gamma^1\gamma^2B^3)\phi_1 \quad (6.2.15)$$

However, the definition (6.2.10) does not allow it, the usual approach and the result just found seem to suggest we may take three other anticommuting elements from (4.5.1) such that (2.2.16) is satisfied. In that fashion, we may write

$$\bar{E}\phi_1=(eA_0-\frac{\hbar^2}{2m}(\nabla-\frac{ie}{\hbar c}\mathbf{A})\cdot(\nabla-\frac{ie}{\hbar c}\mathbf{A})-\frac{e\hbar}{2mc}\sigma\cdot\mathbf{B})\phi_1 \quad (6.2.16)$$

which is the *Pauli equation*, the expected nonrelativistic limit.

Being able to find this result, even though it was done with the usual approach as a strong guideline, leads us to believe that a biquaternion formulation of Dirac's equation may contain all the physical content that is to be expected. We may check this by deriving the spin operators in the biquaternion formalism from (6.2.16). To confirm this idea in a different way, let us look at plane wave solutions of the Dirac equation in the next section.

6.3 Plane wave solutions

Let us look at plane wave solutions of the biquaternion Dirac equation. In the previous section, we chose (4.5.6) as a suitable representation of the Dirac equation to take the nonrelativistic limit. Here we take Conway's biquaternion Dirac equation (4.5.4) to investigate the solutions

$$\chi(x,t) = e^{-\frac{i}{\hbar}(p^0x^0 - \mathbf{p}\cdot\mathbf{x})}u(p) \tag{6.3.1}$$

We will again use the Hamiltonian and the momentum operator from (6.2.1) to arrive at

$$E\chi i_1 - c(p^1 + p^2 + p^3)\chi i_2 + imc^2\chi i_3 = 0 \tag{6.3.2}$$

Note that here we speak of a plane wave corresponding to a particle of positive energy expressed by the sign of E. Repetition of our operators gives

$$(E^2 - c^2((p^1)^2 + (p^2)^2 + (p^3)^2) - m^2c^4)\chi = 0 \tag{6.3.3}$$

which implies the relativistic energy-momentum relation (3.2.1). So, the plane wave (6.3.1) corresponds to a free particle.

When looking at the particle in the rest frame we have

$$\mathbf{p} = 0 \text{ and } E = E_0 = mc^2 \tag{6.3.4}$$

This results in

$$\chi i_1 + i\chi i_3 = e^{-\frac{i}{\hbar}(p^0x^0 - \mathbf{p}\cdot\mathbf{x})}u(p)i_1 + ie^{-\frac{i}{\hbar}(p^0x^0 - \mathbf{p}\cdot\mathbf{x})}u(p)i_3 = 0 \tag{6.3.5}$$

so that we must have

$$u(p)i_1 + iu(p)i_3 = 0 \tag{6.3.6}$$

This relation is satisfied when the function $u(p)$ is

$$u(p) = c_1(1 + ii_2) + c_2(i_1 + ii_3) \tag{6.3.7}$$

for any two complex numbers c_1 and c_2.

Thus, for the positive energy case in the rest frame we have two independent solutions,

$$u_1(p) = 1 + ii_2 \tag{6.3.8}$$

$$u_2(p) = i_1 + ii_3 = i_1(1 + ii_2) = i_1u_1(p) \tag{6.3.9}$$

We seem to have found a geometric interpretation of multiplication by i_1 from the left here: it produces a spin flip.

A plane wave corresponding to negative energy is given by

$$\chi(x,t) = e^{+\frac{i}{\hbar}(p^0x^0 - \mathbf{p}\cdot\mathbf{x})}v(p) \tag{6.3.10}$$

Again, going through the above procedure shows us that there are two indepedent solutions for $v(p)$,

$$v_1(p) = 1 - ii_2 \tag{6.3.11}$$

$$v_2(p) = i_1 - ii_3 = i_1(1 - ii_2) = i_1v_1(p) \tag{6.3.12}$$

There seems to be a pretty straight-forward way to relate the positive and negative energy solutions, namely complex conjugation. In the representation chosen this is indeed the case, but in others the complex unit i does not appear in the solutions. A working relation that signifies a rotation over 180 degrees around the x-axis is given by

$$v_1(p) = 1 - ii_2 = i_1(1 + ii_2)(-i_1) = i_1u_1(p)(-i_1) = u_2(p)(-i_1) \tag{6.3.13}$$

$$v_2(p) = i_1 - ii_3 = i_1(i_1 + ii_3)(-i_1) = i_1u_2(p)(-i_1) = u_1(p)i_1 \tag{6.3.14}$$

We propose that right multiplication by i_1 may be interpreted as going from positive to negative energy solutions. From the used biquaternion Dirac equation (4.5.4) we may see this is obvious. Taking the multiplication by i_1 through all quantities so that it ends up on the right leaves the first term with a relative minus sign.

We are unsure of the role the chosen representation has in the geometric interpretation of left and right multiplication by i_1, however. Perhaps, in choosing a γ^0 in which the units i_2 or i_3 occur will force us to leave our geometric interpretation of i_1. It may be interesting to look into this in further research to establish a possible advantage of the biquaternion formalism.

6.4 Conclusion

We have shown that the biquaternion formulation of Dirac's equation from previous chapters is equivalent to the usual formulation in the sense that it produces the same physics. We have successfully reached the Pauli equation in making the nonrelativistic approximation and we were able to find the plane wave solutions of the Dirac equation quite easily. Note that in this last result we also found that there may be clear geometric interpretations for multiplication with the quaternion units, namely spinflip and a transition from positive to negative energy solutions and vice versa.

Chapter 7

Conclusions

In this thesis, we have aimed to explore the possibility to apply a formalism of (complex) quaternions to the Dirac equation to investigate claims of elegance and intuitiveness of such a formalism and of "new physics".

To this purpose the quaternions and biquaternions have been studied intensively first. Here, we encountered problems in relating four-vectors with quaternions, and therefore, we chose to identify them with biquaternions.

After introduction of the Dirac equation and the associated Dirac algebra, we found that the quaternions were usefull in finding representations of both the Pauli and Dirac algebra. Furthermore, the biquaternions turned out to provide us with one-dimensional representations of the Dirac matrices.

These results we applied to find (complex) quaternion formulations of the Dirac equation. The one-component biquaternion formulation found may certainly be regarded as elegant. Here, the gamma matrices need not be invoked anymore. Also, the description of classical electromagnetism in terms of biquaternions turned out to be particularly ingenious.

In investigating the found biquaternion Dirac equation some nice results were met, namely the possibility to take the nonrelativistic limit and to find plane wave solutions in which there seemed to exist a geometric interpretation of the quaternion units.

Overall, we may say the formalism of biquaternions did well in formulating the Dirac equation and some well-known physical results. We must arrive at the conclusion, however, that reformulating Dirac's equation in terms of biquaternions cannot be called particularly useful. Finding new physical results turns out to be highly unlikely and the formalism has not been able to deepen our understanding of the Dirac equation. Besides, the anticommutativity and the underdevelopedness of the system of quaternions in comparison with the notation usually used in relativistic quantum mechanics were not very helpful, either. Only the geometric interpretations encountered in the last chapter may shed a brighter light on the question of usefulness of the formalism.

Perhaps some of the elegance of application of the algebra of quaternions to the Dirac equation was lost in the extension to complex quaternions. As Dirac remarked, the complex quaternions seem to contain too many components for a simple description of quantities in space-time. Furthermore, it may not have been the right choice to force the Dirac equation in the mould of a biquaternion formalism in the first place, even though it seemed to produce elegant results in formulating Lorentz transformations.

Let us leave the quaternion formalism as a successful description of rotations in three-dimensional space and as a curious way to formulate electrodynamics and Lorentz transformations. Perhaps it once again resurfaces when the well-established notation of relativistic quantum mechanics turns out to be insufficient or when a devotee applies it to yet another field of physics.

Bibliography

[1] W.R. Hamilton. *Lecture on quaternions.* Dublin: Hodges and Smith, etc., 1853.

[2] A.M. Bork. Vectors versus quaternions - The letters in Nature. 34(3):202–211, 1966.

[3] A.W. Conway. The quaternionic form of relativity. *Phil. Mag.*, 24(139):208, 1912.

[4] A.W. Conway. Quaternionic treatment of the relativistic wave equation. *Proc. R. Soc. A*, 162(909):145–154, 1937.

[5] P.A.M. Dirac. Application of quaternions to Lorentz transformations. *Proc. Roy. Irish Acad. A*, 50(16):261–270, 1945.

[6] J.D. Edmonds, Jr. Quaternion quantum theory: new physics or number mysticism? 42(3):220–223, 1974.

[7] S. De Leo. One component Dirac equation. *Int. J. Mod. Phys. A*, 11, 1996.

[8] D. Schuricht and M. Greiter. Contemplations on Dirac's equation in quaternionic coordinates. *Eur. J. Phys.*, 25(6):755–764, 2004.

[9] S. De Leo. Quaternions and Special Relativity. *J. Math. Phys.*, 37(6):2955–2968, 1996.

[10] M. Greiter and D. Schuricht. Imaginary in all directions: an elegant formulation of Special Relativity and classical electrodynamics. *Eur. J. Phys.*, 24(4):397–401, 2003.

[11] J. Lambek. If Hamilton had prevailed: quaternions in physics. *The Mathematical Intelligencer*, 17(4):7–15, 1995.

[12] J.D. Edmonds, Jr. Nature's natural numbers: relativistic quantum theory over the ring of complex quaternions. *Int. J. Theor. Phys.*, 6(3):205–224, 1972.

[13] P.R. Girard. The quaternion group and modern physics. *Eur. J. Phys.*, 5(1):25–32, 1984.

[14] H.F. Jones. *Groups, representations and physics.* Taylor & Francis Group, LLC, 2nd edition, 1998.