# Quaternionic Wavefunction

## Pavel A. Bolokhov

*Theoretical Physics Department, St.Petersburg State University, Ulyanovskaya 1, Peterhof, St.Petersburg, 198504, Russia*

## Abstract

We argue that quaternions form a natural language for the description of quantum-mechanical wavefunctions with spin. We use the quaternionic spinor formalism which is in one-to-one correspondence with the usual spinor language. No unphysical degrees of freedom are admitted, in contrast to the majority of literature on quaternions. In this paper we first build a Dirac Lagrangian in the quaternionic form, derive the Dirac equation and take the non-relativistic limit to find the Schrödinger's equation. We show that the quaternionic formalism is a natural choice to start with, while in the transition to the non-interacting non-relativistic limit the quaternionic description effectively reduces to the regular complex wavefunction language. We provide an easy to use grammar for switching between the ordinary spinor language and the description in terms of quaternions. As an illustration of the broader range of the formalism, we also derive the Maxwell's equation from the quaternionic Lagrangian of Quantum Electrodynamics. In order to derive the equations of motion, we develop the variational calculus appropriate for this formalism.

# 1   Introduction

The list of literature on the rôle of quaternions in physics is so vast that it is hardly possible to enlist it [1], The majority of the literature touches base on the use of quaternions in three-dimensional rotations and Lorentz transformations. Other works include their applications in Electrodynamics and Quantum Mechanics. The attractive feature of quaternions is that whenever a solution is possible in the quaternionic form, it is much more compact than that in the customary form of Lorentz vectors and tensors.

For the majority of theorists, however, quaternions are seen as a somewhat exotic subject, which neither has proven to be exceedingly effective, nor has lead to any new insights or new formalisms.

The other drawback, as it is perceived, is a somewhat "strange" mathematical language, often accompanied by strange results following from it. While from learning Quantum Mechanics and Gauge Theories we are used to non-commutative operators, and non-commutative objects in general, when we see an operator of multiplication "from the right" $|\,\hat{\imath}$ (an example of the so-called "barred" operators) it immediately induces a certain degree of skepsis. That is, mathematically this formalism may be interesting, but physically this seems to be driven away from reality and therefore deemed unnecessary.

In the literature, quaternionic wavefunction is usually introduced *ad hoc*. The wavefunction is named quaternionic just as an attempt of a generalization of Quantum Mechanics, Quantum Electrodynamics, Gravity and so on. Even though such a theory may involve a new form of analysis, new operator formalisms, and other attractive mathematical features, it is not warranted by experiment in any sense. The resulting Dirac or Schrödinger equations often include unobserved degrees of freedom. This is the consequence of the fact that quaternions are multi-component numbers, with a bit too many components than needed for Quantum Mechanics. These aspects of using quaternions in theoretical physics are enough to discourage the interest in the majority of theorists.

Recently, as series of works have been published where a formalism has been developed for constructing the quaternionic analogues of spinors, based on the determination of the so-called maximally totally isotropic subspaces which project the desired minimal ideals from the complex quaternionic algebra $\mathcal{C} \otimes \mathcal{H}$. This construction has a direct physical interpretation in terms of chirality of spinors, which we

actively exploit in this paper. This way we are able to ensure that the content of the theory we are writing does not include any exotic or non-observed degrees of freedom.

Our main goal is to cast a bridge from the regular algebraic language of physics (which is, dominantly based on complex numbers) to the language of quaternions. We argue that quaternions have always been around, and we just neglected to acknowledge them. There is no need or necessity for any new degrees of freedom or new physics to arise. We would like to present a concise dictionary, so that any theorist could connect to and appreciate the quaternionic formalism, which appears to be quite capacious.

The omnipresence of quaternions is easy to observe. We know that Quantum Mechanics is based solely on complex numbers. Complex numbers provide a compact and meaningful way of both formulating and solving quantum-mechanical problems. With some exceptions, it would be very awkward to split the Schrödinger equation into its real and imaginary parts, and then to attempt to solve the resulting system of equations. Quantum mechanical operators of momentum, angular momentum are *inherently* complex. That is, to say, that the wavefunction is complex too.

But as soon as the relativistic effects are included into Quantum Mechanics, it turns out that particles have *spin*. The way to incorporate spin into the wavefunction is just to turn it into a spinor. The spin operator itself is then given by the Pauli matrices. This is where quaternions get involved. The algebra of Pauli matrices is the same as that of quaternionic units, loosely speaking. We argue that there is *no inherent need* to having introduced matrices. Quaternionic units are all that is needed, and *implicitly* the wavefunction in Quantum Mechanics is quaternionic.

So, how does one include spin into the wavefunction by means of quaternions? We proceed in a very conservative way, essentially expanding on the development originally presented in [*thesis*]: we start out from the Dirac Lagrangian, carefully taking into account the spin degrees of freedom, and then *derive* the Schrödinger equation. Spin naturally appears as a consequence of the fact that the wavefunction is quaternionic. Quaternionic derivations, while a bit unusual for some, are simpler than spinor derivations. This way we argue that quaternionic language is natural for Quantum Mechanics.

A well-known peculiar property of physics is that physical objects are actually described by *complex* quaternions, which are also known as complexified quaternions, and not by the regular quaternions. This does not change our main argument,

however. Quaternionic structure allows one pass from the matrix formulation of the theory to the algebraic description. In order to work with spinors one normally has to involve Dirac or Pauli matrices, which often conceal the underlying algebraic structure of the solution. Needless to say, that the equations take a lot more attractive look when such matrices are not involved.

It is not just for these reasons that we believe quaternions play a fundamental rôle in physics. There are hints that quaternions are part of the natural language for the entire Standard Model []. We view this work as one of the steps towards the description of the Standard Model in such a language.

<div align="center">*        *        *</div>

Let us talk about our notations first, while gradually introducing the subject matter. The reader eager to see the physical results may choose to skip to the next section, returning here to clarify the notations when necessary. Here we overview the known facts which are easy to pick up and use, while their proof can be found in the literature. We denote the quaternionic units as

$$\hat{\imath}^2 \;=\; \hat{\jmath}^2 \;=\; \hat{k}^2 \;=\; -1\,, \tag{1.1}$$

and we do not distinguish them from the three-dimensional spatial unit vectors. That is, any three-dimensional vector is a quaternion

$$\vec{a} \;=\; a^1\,\hat{\imath} \;+\; a^2\,\hat{\jmath} \;+\; a^3\,\hat{k}\,. \tag{1.2}$$

Complex quaternions are defined as

$$a \;=\; a^0 \;+\; a^1\hat{\imath} \;+\; a^2\hat{\jmath} \;+\; a^3\hat{k}\,, \tag{1.3}$$ {cq}

where all components

$$a^0 \;=\; b^0 \;+\; i\,b^0\,, \qquad a^1 \;=\; b^1 \;+\; i\,b^1\,, \qquad \ldots, \tag{1.4}$$

are complex numbers. Here $i$ denotes a regular imaginary complex unit, which commutes with $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$.

Let us introduce the conjugation operations. We denote the quaternionic conjugation (q.c.) by $\widetilde{a}$,

$$\widetilde{a} \;=\; a^0 \;-\; a^1\hat{\imath} \;-\; a^2\hat{\jmath} \;-\; a^3\hat{k}\,. \tag{1.5}$$

<div align="center">3</div>

We remember that the quaternionic conjugation switches the order of factors in a product:

$$\widetilde{a\,b} \;=\; \widetilde{b}\,\widetilde{a}\,. \tag{1.6}$$

Since the components are complex numbers, we also have the complex conjugation (c.c.) $a^*$

$$a^* \;=\; (a^0)^* \;+\; (a^1)^*\,\hat{\imath} \;+\; (a^2)^*\,\hat{\jmath} \;+\; (a^3)^*\,\hat{k}\,. \tag{1.7}$$

It is convenient to introduce the composition of these two conjugations, which we call a *hermitean* conjugation (h.c.) (not without a reason),

$$a^\dagger \;=\; (a^0)^* \;-\; (a^1)^*\,\hat{\imath} \;-\; (a^2)^*\,\hat{\jmath} \;-\; (a^3)^*\,\hat{k}\,. \tag{1.8}$$

By itself it does not give anything new, as it is merely a combination of $\tilde{\;}$ and $*$ operations, but it is important to have it, as we will see. Hermitean conjugation also interchanges the order of terms in a product, obviously

$$(a\,b)^\dagger \;=\; b^\dagger\,a^\dagger\,. \tag{1.9}$$

It is this plentitude of conjugations that give richness to the quaternionic language, when applied to physics.

Now we can define a *true* four-dimensional vector

$$v \;=\; v^0 \;+\; i\,(\,v^1\,\hat{\imath} \;+\; v^2\,\hat{\jmath} \;+\; v^3\,\hat{k}\,) \;\equiv\; v^0 \;+\; i\,\vec{v} \tag{1.10} \quad \texttt{\{true-vector\}}$$

in Minkowski space. Note that it is precisely the combination $i\,\vec{v}$ that gives a true vector. Oppositely,

$$v \;=\; i\,v^0 \;+\; \vec{v} \tag{1.11} \quad \texttt{\{pseudo-vector}$$

describes a *pseudo*-scalar and a *pseudo*-vector, with respect to inversion. Together, the two objects (1.10) and (1.11) span the entire space of complex quaternions (1.3). In other words, a generic complex quaternion $a$ can be split into two four-vectors. Their time components will represent a true- and a pseudo-scalar, while their three-dimensional parts will represent a true and an axial vector, correspondingly. Notice, how multiplication by the complex $i$ turns a true vector into an axial vector, and the same for scalars.

With four-dimensional vectors in Minkowski space, one has to be careful to always keep in mind whether a vector is contravariant or covariant. Complex conjugation $*$ turns a contravariant vector $v$ into a covariant vector $v^*$.

$$v^* \;=\; v^0 \;-\; i\,\vec{v}\,. \tag{1.12}$$

4

This actually is equivalent to quaternionic conjugation $v^* = v^\dagger$.

The derivative operator $\partial$

$$\partial = \partial^0 + i\vec{\nabla} \tag{1.13}$$

is *by definition* a covariant vector, while obviously

$$\partial^* = \partial^0 - i\vec{\nabla} \tag{1.14}$$

is a contravariant vector.

To make these identifications more meaningful, let us talk about Lorentz transformations. Lorentz transformations are generated by a purely-imaginary quaternionic parameter

$$\Lambda = \vec{\kappa} + i\vec{\lambda}. \tag{1.15}$$

We are not going to explicitly treat it as a vector, so we are not putting a vector sign on this parameter. Parameter $\vec{\kappa}$ generates three-dimensional rotations, while parameter $\vec{\lambda}$ generates boosts. These are "generators" in the sense that the actual finite transformations are performed by the exponent

$$e^\Lambda. \tag{1.16}$$

It is important to be careful here, as $\vec{\kappa}$ can be interpreted as a three-dimensional rotation only when $\vec{\lambda} = 0$, and the same for $\vec{\lambda}$ — it can be interpreted as a boost only when $\vec{\kappa} = 0$. This is because in general $\vec{\kappa}$ and $\vec{\lambda}$ do not commute, and therefore

$$e^{\vec{\kappa} + i\vec{\lambda}} \neq e^{\vec{\kappa}} \cdot e^{i\vec{\lambda}}, \tag{1.17}$$

where each individual exponent on the right-hand side *is* treated as a rotation and a boost, correspondingly. Notice that since $\Lambda$ is purely imaginary, $\widetilde{\Lambda} = -\Lambda$, and therefore

$$\widetilde{(e^\Lambda)} = e^{-\Lambda}. \tag{1.18}$$

A contravariant vector $v$ transforms under $\Lambda$ as

$$v \rightarrow e^\Lambda v e^{\Lambda^\dagger}. \tag{1.19} \quad \{\texttt{contra}\}$$

Any covariant vector then should transform the same way that $v^*$ does:

$$v^* \rightarrow e^{\Lambda^*} v^* e^{\widetilde{\Lambda}}. \tag{1.20} \quad \{\texttt{co}\}$$

As a special case of these, a three-dimensional vector $\vec{v}$ rotates as

$$\vec{v} \;\; \rightarrow \;\; e^{\vec{\kappa}}\,\vec{v}\,e^{-\vec{\kappa}}\,. \qquad (1.21) \quad \{\texttt{rotation}\}$$

This concludes our basic discussion of vectors for now.

Another way a complex quaternion (1.3) can be split up, is by separating its zeroth component $a^0 \;\equiv\; \phi$ and its vector part $\vec{a}$. The zeroth component $\phi$ is a complex number whose real and imaginary parts are identified as a true and axial scalar fields, correspondingly. The remaining vector part is then identified as a field strength:

$$\vec{F} \;\; = \;\; \vec{B} \;+\; i\,\vec{E}\,. \qquad (1.22) \quad \{\texttt{fs}\}$$

Notice that this agrees with the identifications of (1.10) and (1.11) as polar and axial vectors. While both $\vec{B}$ and $\vec{E}$ are vectors in the three-dimensional sense, one does not view the object (1.22) as a four-dimensional vector in any way. This is an entirely different split up of a complex quaternion.

Both $\phi$ and $\vec{F}$ transform the same way under Lorentz transformations:

$$p \;\; \rightarrow \;\; e^{\Lambda}\,p\,e^{\widetilde{\Lambda}}\,. \qquad (1.23)$$

For the scalar $\phi$ this obviously does not do anything, since $\phi$ is just a complex number:

$$e^{\Lambda}\,\phi\,e^{\widetilde{\Lambda}} \;\; = \;\; \phi\,e^{\Lambda}\,e^{\widetilde{\Lambda}} \;\; = \;\; \phi\,,$$

so it is indeed a scalar. While for the field strength, this dictates that

$$\vec{F} \;\; \rightarrow \;\; e^{\Lambda}\,\vec{F}\,e^{\widetilde{\Lambda}}\,. \qquad (1.24)$$

Note that for the case of pure rotations, when $\Lambda \;=\; \vec{\kappa}$, this agrees with Eq. (1.21), as $\widetilde{\Lambda} \;=\; -\Lambda \;=\; -\vec{\kappa}$. That is, both the electric and magnetic fields rotate as three-vectors.

Finally, we introduce spinors. Here we give a very brief overview of spinors necessary for Section 2, while a more detailed discussion is postponed until Appendix A. We begin with stating that the space of complex quaternions $\mathcal{C} \otimes \mathcal{H}$ can be split in two halves in a yet another way, namely using chirality projectors, $P_L$ and $P_R$. This construction is based on identifying the projectors corresponding to the maximally totally isotropic subspaces of the quaternion algebra [$thesis$]. In the case of $\mathcal{C} \otimes \mathcal{H}$ such spaces are one-dimensional each, and are given by the projectors $P_L$ and $P_R$.

6

A *left-handed* spinor is defined as an arbitrary complex quaternion multiplied by a projector $P_L$ on the right:

$$\psi_L \;=\; a\,P_L. \tag{1.25}$$

Here $P_L$ is a complex quaternion

$$P_L \;=\; \frac{1 + i\hat{k}}{2}, \tag{1.26}$$

and we call it a projector because

$$P_L^2 \;=\; P_L.$$

All accompanying details of these definitions can be found in Appendix A. For an ordinary reader $\psi_L$ should be precisely viewed as a left-handed chiral spinor. A convenient basis for $\psi_L$ is formed by elements $P_L$ and $\hat{\jmath}\,P_L$, which are geometrically orthogonal to each other:

$$\psi_L \;=\; \xi_L\,P_L \;+\; \chi_L\,\hat{\jmath}\,P_L. \tag{1.27} \quad \{\texttt{lbasis}\}$$

Complex numbers $\xi_L$ and $\chi_L$ are precisely the "spin-up" and "spin-down" components of $\psi_L$ viewed as a Weyl spinor:

$$\psi_L \;=\; \begin{pmatrix} \xi_L \\ \chi_L \end{pmatrix}. \tag{1.28}$$

Left-handed spinors span one half of the complex quaternion space, while the other half is spanned by the right-handed spinors:

$$\psi_R \;=\; a\,P_R, \tag{1.29}$$

where

$$P_R \;=\; \frac{1 - i\hat{k}}{2}. \tag{1.30}$$

These two halves are related by complex conjugation $*$, and the projectors are related as

$$P_R \;=\; P_L^*, \qquad\qquad P_L \;+\; P_R \;=\; 1. \tag{1.31}$$

The basis for $\psi_R$ is similarly given by $P_R$ and $\hat{j}\,P_R$, but the components are identified slightly differently

$$\psi_R \;=\; -\,\xi_R\,\hat{j}\,P_R \;+\; \chi_R\,P_R\,. \tag{1.32} \quad \texttt{\{rbasis\}}$$

Defined like so,

$$\psi_R \;=\; \begin{pmatrix} \xi_R \\ \chi_R \end{pmatrix} \tag{1.33}$$

is precisely identified as a right-handed Weyl spinor. The basis (1.27) and especially so (1.32) may seem a bit awkward, but they are convenient for doing algebra. The fact that the projectors are part of the bases allows us to perform various manipulations and conjugations on spinors quite effectively.

Under Lorentz transformations, left spinors by definition transform as

$$\psi_L \;\rightarrow\; e^{\Lambda}\,\psi_L\,, \tag{1.34}$$

while the right-handed ones transform in the conjugate representation

$$\psi_R \;\rightarrow\; e^{\Lambda^*}\,\psi_R\,. \tag{1.35}$$

It is interesting to observe and compare how spinors and three-dimensional vectors transform under three dimensional rotations. Consider a rotation around axis $\hat{a}$ $(\hat{a}^2 \;=\; -\,1)$ through angle $\alpha$. Vector $\vec{v}$ rotates as

$$\vec{v} \;\rightarrow\; e^{\alpha\,\hat{a}/2}\,\vec{v}\,e^{-\alpha\,\hat{a}/2}\,, \tag{1.36}$$

while spinors transform as

$$\psi_{L,R} \;\rightarrow\; e^{\alpha\,\hat{a}/2}\,\psi_{L,R}\,. \tag{1.37}$$

In particular, if we perform a rotation through $2\pi$, then $e^{\pm 2\pi\,\hat{a}/2} \;=\; -\,1$, and a vector is unchanged, while a spinor changes its sign, as it should be. Of course, this is because of the factor of $1/2$ in the exponent (the famous "Rodrigues' two"), which is just the reflection of the fact that SU(2) is a double cover of the rotation group SO(3).

The action of discrete symmetries on spinors is presented in Appendix A.1.

# 2 Dirac equation

Our goal in this section is to construct the Lagrangian for the electron in electromagnetic field, in quaternionic form, and derive the equation of motion — the Dirac equation.

Let us begin with a massless particle, in the absence of the electromagnetic field. The appropriate Lagrangian was given in [*thesis*],

$$\mathcal{L}_{\text{massless}} \;\; = \;\; \psi_L^\dagger \, i\partial \, \psi_L \;\; + \;\; \psi_R^\dagger \, i\, \partial^* \, \psi_R \;\; + \;\; \text{c.c.} \tag{2.1} \quad \text{\{L-massless\}}$$

Note that Lorentz invariance is manifest here, because

$$\psi_L^\dagger \;\; \rightarrow \;\; \psi_L^\dagger \, e^{\Lambda^\dagger} \;\; = \;\; \psi_L^\dagger \, e^{-\Lambda^*}, \quad \partial \;\; \rightarrow \;\; e^{\Lambda^*} \, \partial \, e^{-\Lambda}, \quad \psi_L \;\; \rightarrow \;\; e^{\Lambda} \, \psi_L, \tag{2.2}$$

and

$$\psi_R^\dagger \;\; \rightarrow \;\; \psi_R^\dagger \, e^{\widetilde{\Lambda}} \;\; = \;\; \psi_R^\dagger \, e^{-\Lambda}, \quad \partial^* \;\; \rightarrow \;\; e^{\Lambda} \, \partial^* \, e^{-\Lambda^*}, \quad \psi_R \;\; \rightarrow \;\; e^{\Lambda^*} \, \psi_R. \tag{2.3}$$

We need to make an important remark about conjugating products of spinors, due to the fact that spinors are Grassmann variables. By definition, complex conjugation of two Grassmann variable interchanges their order,

$$(\zeta \, \eta)^* \;\; = \;\; \eta^* \, \zeta^*, \qquad \text{for Grassmann numbers.} \tag{2.4}$$

If we take two complex quaternions $\xi$ and $\chi$, which are fermions at the same time, complex conjugation cannot change their order, because of their quaternionic content. In that case, the order is preserved, but an extra minus sign appears,

$$(\xi \, \chi)^* \;\; = \;\; - \, \xi^* \, \chi^*, \qquad \text{for fermionic complex quaternions.} \tag{2.5} \quad \text{\{ferm-cc\}}$$

If we take a quaternionic conjugate $\widetilde{\xi \chi}$, on the other hand, the conjugation will attempt to change their order precisely because of the quaternionic content. Note that the quaternionic algebra *requires* us to interchange the factors, or the result will simply be incorrect. But now because the spinors are fermions, and we are *not* performing a complex conjugation, we get an extra minus sign

$$\widetilde{\xi \chi} \;\; = \;\; - \, \widetilde{\chi} \, \widetilde{\xi} \qquad \text{for quaternionic spinors.} \tag{2.6} \quad \text{\{ferm-qc\}}$$

9

The only kind of conjugation which does not produce a negative sign is hermitean conjugation — this combination changes the order of the spinors in agreement with both complex and quaternionic conjugations,

$$(\xi\,\chi)^\dagger \;=\; \chi^\dagger\,\xi^\dagger\,. \qquad (2.7) \quad \texttt{\{ferm-hc\}}$$

Lagrangian (2.1) is manifestly chiral, in that it consists of two separate terms for the left-handed and right-handed fermions. This might be considered as a disadvantage compared to the usual matrix-based representation, where both the mass term and the kinetic term are united. In quaternionic language, we keep the chiralities in Eq. (2.1) separate here. As we will be dealing with the quantum-mechanics limit, explicit chiral structure is more advantageous in this study. Furthermore, the Standard Model, the way to which this work is meant to point, is chiral in its nature. It is, however, possible to unite the terms in Eq. (2.1). To do that, we need to appeal to the so-called $m_c$-action [*thesis*] that the derivatives in (2.1) implement. In plain terms, the kinetic term involves the projectors acting on the right,

$$\mathcal{L}_{\mathrm{massless}} \;=\; \psi_D^\dagger\, i\left(\partial\,|\,P_L \;+\; \partial^*\,|\,P_R\right)\psi_D \;+\; \mathrm{c.c.} \qquad (2.8) \quad \texttt{\{L-united\}}$$

Expressing the kinetic term in this form will be useful for promoting the theory to the quantum stage and developing the Feynman rules. This will be the subject of future work.

Let us discuss gauge transformations now, as we need the electron to interact with the electromagnetic field. Gauge transformations just rotate the overall complex phase of a spinor, and so they are defined similarly both for right- and left-handed spinors,

$$\psi_L \;\rightarrow\; e^{i\varphi}\,\psi_L\,, \qquad\qquad \psi_R \;\rightarrow\; e^{i\varphi}\,\psi_R\,. \qquad (2.9)$$

Notice that the quaternionic conjugates also transform the same way,

$$\widetilde{\psi}_L \;\rightarrow\; \widetilde{\psi}_L\,e^{i\varphi}\,, \qquad\qquad \widetilde{\psi}_R \;\rightarrow\; \widetilde{\psi}_R\,e^{i\varphi}\,. \qquad (2.10)$$

Although $e^{i\varphi}$ certainly commutes with $\widetilde{\psi}_{L,R}$, here for convenience we wrote it on the right of the latter. Both the complex conjugates and hermitean conjugates will have the opposite charge,

$$\psi_L^* \;\rightarrow\; \psi_L^*\,e^{-i\varphi}\,, \qquad\qquad \psi_R^* \;\rightarrow\; \psi_R^*\,e^{-i\varphi}\,,$$

$$\psi_L^\dagger \;\rightarrow\; \psi_L^\dagger\,e^{-i\varphi}\,, \qquad\qquad \psi_R^\dagger \;\rightarrow\; \psi_R^\dagger\,e^{-i\varphi}\,. \qquad (2.11)$$

10

In order to make this transformation local, we define the long derivative,

$$\mathcal{D} \;=\; \partial \;-\; i\,A^{*}\,. \tag{2.12}$$

The reason that we have to put $A^{*}$ here instead of just $A$ is because $\partial$ and $\mathcal{D}$ are covariant vectors. This is just the reflection of the fact that $A^{\mu}$ enters the long derivative with the lower index $\mu$:

$$\mathcal{D}_{\mu} \;=\; \partial_{\mu} \;-\; i\,A_{\mu}\,. \tag{2.13}$$

This long derivative then transforms as

$$\mathcal{D} \;\rightarrow\; e^{i\varphi}\,\mathcal{D}\,e^{-i\varphi}\,, \tag{2.14}$$

meaning that, as usual,

$$A_{\mu} \;\rightarrow\; A_{\mu} \;+\; \partial_{\mu}\,\varphi\,. \tag{2.15}$$

This allows $\mathcal{D}$ to act on $\psi_L$, so that $\mathcal{D}\,\psi_L$ is again in the fundamental representation of U(1). Now, although $\psi_R$ has the same charge as $\psi_L$, we cannot act on it with the same derivative, because the product $\mathcal{D}\,\psi_R$ will not transform under the Lorentz transformations properly. Remind, that the right-handed and left-handed spinor spaces are in fact related by complex conjugation. This was the reason that we wrote $\partial^{*}$ in the Lagrangian in Eq. (2.1). One would think that by analogy we should act on $\psi_R$ with $\mathcal{D}^{*}$ — but that would also be a mistake because it would imply that $\psi_R$ has the opposite charge. In reality, it is the quaternionic conjugate $\widetilde{\mathcal{D}}$ that should be put into the Lagrangian. This conjugation does not change the sign of the electric charge. Overall, the Lagrangian now looks as,

$$\mathcal{L}_{\text{massless}} \;=\; \psi_L^{\dagger}\,i\mathcal{D}\,\psi_L \;+\; \psi_R^{\dagger}\,i\widetilde{\mathcal{D}}\,\psi_R \;+\; \text{c.c.} \tag{2.16} \quad \{\texttt{L-gauge}\}$$

Before proceeding, let us emphasize the remarkable feature of the Lagrangians (2.1) and (2.16), which we could have done earlier: the absence of $\gamma$-matrices (or $\sigma$-matrices for that matter). The only residue of the matrix structure of the Dirac's Lagrangian is residing in the fact that the Lagrangians have two chiral terms, instead of just one. We will return to this below.

Now we add the mass term. Since this has to be the Dirac mass, it has to flip chirality. The only form that correctly reproduces the mass term has the form $m\,\hat{\jmath}\,\psi_L^{\dagger}\,\psi_R$,

$$m\,\hat{\jmath}\,\psi_L^{\dagger}\,\psi_R \;+\; \text{q.c.} \;+\; \text{c.c} \;=\; m\left( \hat{\jmath}\,\psi_L^{\dagger}\,\psi_R \;+\; \widetilde{\psi}_R\,\psi_L^{*}\,\hat{\jmath} \right) \;+\; \text{c.c.} \tag{2.17}$$

11

Here the signs q.c. and c.c. imply adding the appropriate conjugate of everything that resides to the left of the respective sign. We will discuss the occurrence of $\hat{j}$ in this expression in detail in Appendix A. This form of the mass term seems awkward, and we will be able to get rid of it soon, after we discover its meaning. For now we just note that without it the expression would vanish, *i.e.* $\psi_L^\dagger \psi_R$ has no real part. Let us also note that this factor of $\hat{j}$ can be moved to the right at our convenience:

$$m\,\psi_L^\dagger\,\psi_R\,\hat{j} \ \ + \ \ \text{q.c.} \ \ + \ \ \text{c.c.} \tag{2.18} \quad \{\texttt{m-term}\}$$

This follows from the general cyclic property of quaternion product under the "q.c." sign, analogous to the cyclicity of trace of matrices[1] —

$$a\,b\,c \ \ + \ \ \text{q.c.} \ \ = \ \ b\,c\,a \ \ + \ \ \text{q.c.} \tag{2.19}$$

The proof is simple — the real part of a product of *two* quaternions cannot depend on their order. From this follows the cyclicity. The fact that we are dealing with complexified quaternions cannot change this property.

As we will see in Appendix A, in the product $\psi_R\,\hat{j}$ the factor $\hat{j}$ "elevates" the right-handed spinor $\psi_R$ to the left-handed space (we have used the term "elevates" because conventionally a left-handed spinor is written above right-handed spinor inside the column of a Dirac spinor). Importantly, the factor $\hat{j}$ does not change the spinor's representation (it is obviously still transformed via multiplication by $e^{\Lambda^*}$ on the left). Instead, the spinor just becomes expandable in the left-handed basis (1.27). We will use this when we define the *standard* representation for spinors below.

Written in terms of the components, the mass term (2.18) gives

$$\psi_L^\dagger\,\psi_R\,\hat{j} \ \ + \ \ \text{q.c.} \ \ + \ \ \text{c.c.} \ \ = \ \ \xi_L^*\,\xi_R \ + \ \chi_L^*\,\chi_R \ + \ \xi_R^*\,\xi_L \ + \ \chi_R^*\,\chi_L\,, \tag{2.20}$$

as it should be for the Dirac mass term.

We also mention here the mass term (2.18) can be written in an alternative form by replacing $\hat{j}$ with $i\hat{i}$,

$$i\hat{i}\,m\,\psi_L^\dagger\,\psi_R \ \ + \ \ \text{q.c.} \ \ + \ \ \text{c.c.} \tag{2.21} \quad \{\texttt{m-term-alt}\}$$

Which form to use is a matter of convenience.

---

[1]In fact, if one chooses to represent quaternions via Pauli matrices, the real part of a quaternion exactly corresponds to the trace of its matrix representation.

## 2.1 Dirac Lagrangian

Now we can derive the Dirac equation. Our starting point is the full Lagrangian

$$\mathcal{L}_{\text{Dirac}} \;=\; \psi_L^\dagger \, i\mathcal{D} \, \psi_L \;+\; \psi_R^\dagger \, i\widetilde{\mathcal{D}} \, \psi_R \;-\; \left( m\,\psi_L^\dagger \, \psi_R \, \hat{j} \;+\; \text{q.c.} \right) \;+\; \text{c.c.}, \quad (2.22) \quad \text{\{L-massive\}}$$

where we remember that the mass term has to actually enter with a negative sign, and we consider the electromagnetic field to be fixed. Having multiple terms, this Lagrangian can also be compacted into a form similar to (2.8). We believe, however, that this should be automatically achieved in the context of the electroweak theory. For our current purposes, once again, it is important to establish the contact with Quantum Mechanics, and we prefer to keep the theory in the chiral form (2.22).

To derive the equations of motion, we vary the Lagrangian (2.22) with respect to $\psi_L^\dagger$ and $\psi_R^\dagger$. The reader may wonder at this point – how are we going to differentiate this Lagrangian with respect to quaternions, let alone complexified and fermionic? We postpone the formal answer to this question until Appendix B. For now, we can just act intuitively, at least when differentiating with respect to $\psi_L^\dagger$. Indeed, in the terms where it is present in Eq. (2.22), it is sitting on the left, and so the naïve left derivative gives

$$i\mathcal{D}\,\psi_L \;-\; m\,\psi_R\,\hat{j} \;=\; 0\,. \quad\quad (2.23) \quad \text{\{ldirac\}}$$

Loosely speaking, for quaternions $\partial\widetilde{q}/\partial q \;\neq\; 0$ — unlike for complex numbers, for which $\partial\overline{z}/\partial z \;=\; 0$. So it seems like there should be more terms on the left-hand side of Eq. (2.23). Why we can act so naïvely, and why the other terms do not contribute is, again, explained in Appendix B. Here we provide an alternative and a more transparent justification, as follows. Let us for a moment pretend that we are dealing with a "quaternionic" Lagrangian

$$\psi_L^\dagger \, i\mathcal{D} \, \psi_L \;+\; m\,\psi_L^\dagger \, \psi_R \, \hat{j} \quad\quad (2.24) \quad \text{\{L-quat\}}$$

from which we will only need the real part. If we find the extremum of this Lagrangian, it will also extremize the real part of the latter. But the extremum of (2.24) is exactly given by Eq. (2.23).

In order to vary with respect to $\psi_R^\dagger$, we just re-write the mass term as

$$\mathcal{L}_{\text{Dirac}} \;\supset\; \left( m\,\psi_R^\dagger \, \psi_L \, \hat{j} \;+\; \text{q.c.} \right) \;+\; \text{c.c.} \quad\quad (2.25)$$

13

We did not introduce anything new, as the first term here was just hidden inside the "q.c." and "c.c." in Eq. (2.22). Notice that this term enters with a *positive* sign. Now we can vary the Lagrangian with respect to $\psi_R^\dagger$, finding

$$i\,\widetilde{\mathcal{D}}\,\psi_R \;+\; m\,\psi_L\,\hat{\jmath} \;=\; 0\,. \qquad (2.26) \quad \texttt{\{rdirac\}}$$

The two expressions (2.23) and (2.26) are the *quaternionic Dirac equations.*

## 2.2   Non-relativistic limit

Now let us derive the Schrödinger equation. We do it in the classical way $[BLP]$, by taking the large-mass limit. First, we open up the long derivatives in Eqs. (2.23), (2.26),

$$i\,\partial_0\,\psi_L \;+\; A_0\,\psi_L \;-\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_L \;-\; m\,\psi_R\,\hat{\jmath} \;=\; 0\,,$$

$$i\,\partial_0\,\psi_R \;+\; A_0\,\psi_R \;+\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_R \;+\; m\,\psi_L\,\hat{\jmath} \;=\; 0\,. \qquad (2.27) \quad \texttt{\{dirac-split\}}$$

We can make one consistency check. For a free particle, the time derivative gives the energy, $i\,\partial_0 \;\to\; E$. If at the same time, it has zero momentum, then $E \;=\; m$, and we find

$$m\,\psi_L \;-\; m\,\psi_R\,\hat{\jmath} \;=\; 0 \qquad\qquad \Longrightarrow \qquad \psi_L \;=\; +\,\psi_R\,\hat{\jmath}\,,$$

$$m\,\psi_R \;+\; m\,\psi_L\,\hat{\jmath} \;=\; 0 \qquad\qquad \Longrightarrow \qquad \psi_R \;=\; -\,\psi_L\,\hat{\jmath}\,. \qquad (2.28) \quad \texttt{\{rest\}}$$

The two equations are consistent. What does the equality $\psi_L \;=\; \psi_R\,\hat{\jmath}$ mean? As we already mentioned, and as we shall see in Appendix A, multiplying by $\hat{\jmath}$ on the right *rotates* the components of $\psi_L$ and $\psi_R$ into each other. That is, in Dirac spinor language, for a spinor

$$\psi_{\text{chiral}} \;=\; \begin{pmatrix} \psi_L^\alpha \\ \psi_{R\dot\alpha} \end{pmatrix} \qquad\qquad (2.29)$$

this operation turns

$$\psi_L^\alpha \;\to\; -\,\psi_{R\dot\alpha}\,, \qquad\qquad\qquad \psi_{R\dot\alpha} \;\to\; \psi_L^\alpha\,. \qquad (2.30)$$

Equation (2.28) then simply implies that the two spinors $\psi_L^\alpha$ and $\psi_{R\dot\alpha}$ are equal. This is correct, as a particle at rest is described by a single two-component spinor.

14

This is too crude a limit, we want to keep the right-handed and left-handed spinors different to the order $O(1/m)$, and momentum non-zero.

To do that, we, essentially, switch to the standard representation. Let us multiply the second equation in (2.27) by $\hat{j}$ on the right:

$$i\,\partial_0\,\psi_L \;+\; A_0\,\psi_L \;-\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_L \;-\; m\,\psi_R\,\hat{j} \;\;=\;\; 0\,,$$

$$i\,\partial_0\,\psi_R\,\hat{j} \;+\; A_0\,\psi_R\,\hat{j} \;+\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_R\,\hat{j} \;-\; m\,\psi_L \;\;=\;\; 0\,. \qquad (2.31) \quad \{\texttt{dirac-j}\}$$

It is natural now to add and subtract these two equations, and introduce the notations

$$\psi_+ \;\equiv\; \zeta \;=\; \frac{\psi_L + \psi_R\hat{j}}{\sqrt{2}}\,, \qquad \psi_- \;\equiv\; \eta \;=\; \frac{\psi_L - \psi_R\hat{j}}{\sqrt{2}}\,. \qquad (2.32) \quad \{\texttt{standard}\}$$

As discussed in Appendix A, these are precisely the upper and lower components of a Dirac spinor in the standard representation:

$$\psi_{\text{standard}} \;=\; \begin{pmatrix} \zeta \\ \eta \end{pmatrix}\,. \qquad (2.33)$$

In this section, we prefer to call them $\psi_\pm$. For a particle at rest, $\psi_- = 0$ as it should be. Equations (2.31) now read

$$i\,\partial_0\,\psi_+ \;+\; A_0\,\psi_+ \;-\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_- \;\;=\;\; m\,\psi_+\,.$$

$$i\,\partial_0\,\psi_- \;+\; A_0\,\psi_- \;-\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_+ \;\;=\;\; -\,m\,\psi_-\,. \qquad (2.34) \quad \{\texttt{dirac-standar}$$

We now get rid of the mass term in the first equation in (2.34) via the redefinition

$$\psi_\pm \;\to\; e^{-imt}\,\psi_\pm\,.$$

The mass term is now gone in that equation:

$$i\,\partial_0\,\psi_+ \;=\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_- \;-\; A_0\,\psi_+\,, \qquad (2.35)$$

while in the second equation the mass term doubles. Taking the large-mass limit we find

$$2m\,\psi_- \;=\; (\vec{\nabla} \;+\; i\,\vec{A})\,\psi_+\,,$$

and so

$$i\,\partial_0\,\psi_+ \;=\; \frac{1}{2m}\,(\vec{\nabla} \;+\; i\,\vec{A})^2\,\psi_+ \;-\; A_0\,\psi_+\,. \qquad (2.36)$$

15

The final step now is to transform the quaternionic square

$$(\vec{\nabla} + i\,\vec{A})^2$$

into the usual vector operators — scalar and vector products. Denoting the scalar-product square of a vector by figure brackets { }, one finds

$$(\vec{\nabla} + i\,\vec{A})^2 \;=\; -\,\{\vec{\nabla} + i\,\vec{A}\}^2 \;+\; i\,\vec{B}\,,$$

in the operator form. We thus get

$$i\,\partial_0\,\psi_+ \;=\; -\,\frac{1}{2m}\,\{\vec{\nabla} + i\,\vec{A}\}^2\,\psi_+ \;+\; \frac{i\,\vec{B}}{2m}\,\psi_+ \;-\; A_0\,\psi_+\,. \qquad (2.37) \quad \{\texttt{schrod}\}$$

Rescaling $A^\mu \to e\,A^\mu$, and switching to the Gaussian units, we arrive to

$$i\,\hbar\,\partial_t\,\psi_+ \;=\; \left(\frac{1}{2m}\,\{\vec{p} + \frac{e}{c}\,\vec{A}\}^2 \;-\; e\,A_0 \;+\; \frac{ie\hbar}{2mc}\,\vec{B}\right)\psi_+\,, \qquad (2.38) \quad \{\texttt{pauli}\}$$

which is nothing but the *Pauli equation*. Notice once again that the $\sigma$-matrices are absent from equations (2.37) and (2.38). Spin is the intrinsic property of the wavefunction.

The factor of $i$ in the magnetic moment term in Eq. (2.38) has a double rôle. First, it is needed by the correspondence

$$\vec{\sigma} \;\to\; i\,\hat{\imath}\,,\;\; i\,\hat{\jmath}\,,\;\; i\,\hat{k}\,,$$

because the $\sigma$-matrices are hermitean, and the quaternionic units are not. Second (and actually the same), by multiplying the magnetic field $\vec{B}$, it turns it from an axial vector into a polar vector $i\,\vec{B}$. In the regular spinor formalism, one yet has to argue that the product $(\vec{\sigma}\cdot\vec{B})$ yields a true scalar when acting on $\psi$.

It is precisely the form of equation (2.38) that establishes both the quaternionic form of the Schrödinger equation and the fact that the Dirac Lagrangian (2.22) is correct. Equation (2.38) involves a single two-component spinor $\psi_+$. As a complex quaternion, $\psi_+$ only occupies a half of the quaternionic space, and in the most general form it can be written as

$$\psi_+ \;=\; \psi\,P_L\,.$$

In fact this restriction can always be imposed in the end. The Schrödinger equation (2.38) can be solved for an arbitrary $\psi$, and then projector $P_L$ can be applied. The

16

fact that the set of left-handed spinors forms an ideal guarantees that the projection will be a solution too.

The right-handed projection $\psi\, P_R$ will also be a solution, of course. It may either represent the same solution, or a different one. It can always be "moved" to the left-handed space by multiplying by $\hat{\jmath}$ on the right.

The magnetic moment term in Eq. (2.38) is the only signature of quaternions in that equation. Without it, the equation looks exactly like the ordinary Schrödinger equation,

$$i\,\hbar\,\partial_t\,\psi \;=\; \left(\frac{p^2}{2m}\,+\,V\right)\psi\,, \tag{2.39} \quad \{\texttt{simple-schrod}$$

where we have dropped the gauge field for simplicity. For a complexified quaternion $\psi$, this equation falls apart into four identical complex equations. So the non-relativistic Quantum Mechanics essentially only needs complex numbers. We stress again that it is the spin of the particle that calls for the quaternionic appearance of the wavefunction.

## 2.3   Current

Let us now find an expression for the electromagnetic current. It can be easily read off the Lagrangian. Indeed, the current by definition is whatever the gauge field couples to,

$$\mathcal{L} \;\supset\; j_\mu\, A^\mu \;=\; \frac{j\,\widetilde{A}\,+\,A\,\widetilde{\jmath}}{2} \;=\; \frac{\widetilde{A}\,j\,+\,\widetilde{\jmath}\,A}{2}\,. \tag{2.40} \quad \{\texttt{L-current}\}$$

We can differentiate (*i.e.* vary) the latter quaternionic expression using equations (B.45) and (B.46). The last fraction in Eq. (2.40) suggests differentiating with respect to $\widetilde{A}$ in order to get current $j$,

$$\partial_{\widetilde{A}}\mathcal{L} \;\supset\; \frac{4\,j\,+\,(-2\,\widetilde{\widetilde{\jmath}})}{2} \;=\; j\,. \tag{2.41}$$

We could now apply the same variational derivative to Lagrangian (2.22). Only this is not necessary — it is enough to re-write the interacting part of (2.22) in the form (2.40) using the cyclicity property,

$$\mathcal{L} \;\supset\; \psi_L^\dagger\,\widetilde{A}\,\psi_L \;+\; \psi_R^\dagger\,A\,\psi_R \;+\; \widetilde{\psi}_L\,A\,\psi_L^* \;-\; \widetilde{\psi}_R\,\widetilde{A}\,\psi_R^*$$

$$=\; -\,\widetilde{A}\,\psi_L\,\psi_L^\dagger \;+\; \widetilde{A}\,\psi_R^*\,\widetilde{\psi}_R \;+\; \text{q.c.}, \tag{2.42}$$

17

from which we conclude that

$$j \;=\; - \, 2 \left( \psi_L \, \psi_L^\dagger \;-\; \psi_R^* \, \widetilde{\psi}_R \right) . \qquad (2.43) \quad \texttt{\{current\}}$$

We immediately observe that this expression transforms correctly under gauge transformations,

$$j \;\to\; e^{i\varphi} \, j \, e^{-i\varphi} \;=\; j \,, \qquad (2.44)$$

and under Lorentz transformations,

$$j \;\to\; e^\Lambda \, j \, e^{\Lambda^\dagger} . \qquad (2.45)$$

Written explicitly in components, expression (2.43) gives,

$$
\begin{aligned}
j \;=\;& \left( \xi_L^* \, \xi_L \;+\; \chi_L^* \, \chi_L \right) \;+\; \left( \xi_R^* \, \xi_R \;+\; \chi_R^* \, \chi_R \right) \;+ \\
&+\; i\hat{\imath} \left( \left( \xi_L^* \, \chi_L \;+\; \chi_L^* \, \xi_L \right) \;-\; \left( \xi_R^* \, \chi_R \;+\; \chi_R^* \, \xi_R \right) \right) \\
&+\; i\hat{\jmath} \left( i \left( \chi_L^* \, \xi_L \;-\; \xi_L^* \, \chi_L \right) \;-\; i \left( \chi_R^* \, \xi_R \;-\; \xi_R^* \, \chi_R \right) \right) \\
&+\; i\hat{k} \left( \left( \xi_L^* \, \xi_L \;-\; \chi_L^* \, \chi_L \right) \;-\; \left( \xi_R^* \, \xi_R \;-\; \chi_R^* \, \chi_R \right) \right) ,
\end{aligned}
\qquad (2.46)
$$

as it should be for the current written in the spinor representation.

Finally, we can re-write current (2.43) using the standard representation (2.32),

$$j \;=\; - \left( \zeta \, \zeta^\dagger \;+\; \eta \, \eta^\dagger \;+\; \text{c.c} \right) \;-\; \left( \zeta \, \eta^\dagger \;+\; \eta \, \zeta^\dagger \;-\; \text{c.c} \right) . \qquad (2.47)$$

Here the first bracket gives the real part — the $j^0$ component, while the second bracket gives the (imaginary) vector part.

# 3 Maxwell's equation

Although not the main subject of our discussion, we can add the kinetic term for the gauge field, and derive the resulting Maxwell's equation. This section mainly serves the purpose of illustration of differentiating with respect to quaternionic vectors. One has to be particularly careful about the targets of differentiation of various involved derivatives.

The gauge field strength is easily found to be

$$\Phi \;=\; \vec{B} \;+\; i\vec{E} \;=\; -\,\frac{\partial A - \widetilde{A}\,\widetilde{\partial}}{2}\,. \qquad (3.48)$$

This expression becomes particularly simple in the Lorentz gauge:

$$\Phi_{\text{Lorentz}} \;=\; -\,\partial A\,, \qquad (3.49)$$

although we will not use it for deriving the equations of motion.

The relevant part of the Lagrangian looks as,

$$\mathcal{L}_{\text{Maxwell}} \;=\; \frac{\tfrac{1}{2}\,\Phi^2 \;+\; \widetilde{A}\,j}{4} \;+\; \text{q.c.} \;+\; \text{c.c.} \qquad (3.50)$$

As we know from Section 2.3, varying the current part of the Lagrangian with respect to $\widetilde{A}$ expectably gives us the current,

$$\partial_{\widetilde{A}}\,\mathcal{L}_{\text{current}} \;=\; \partial_{\widetilde{A}}\,\frac{\widetilde{A}\,j + \widetilde{j}\,A}{2} \;=\; j\,. \qquad (3.51)$$

Now we need to vary the gauge part of the Lagrangian,

$$\mathcal{L}_{\text{gauge}} \;=\; \frac{1}{4}\,\Phi^2 \;+\; \text{c.c.} \qquad (3.52)$$

Explicitly,

$$\partial_{\widetilde{A}}\,(\partial A \;-\; \widetilde{A}\,\widetilde{\partial})\,(\partial A \;-\; \widetilde{A}\,\widetilde{\partial}) \;=\;$$

$$\underline{\underline{\partial}}_{\widetilde{A}}\,(\partial\underline{\underline{A}} \;-\; \underline{\underline{\widetilde{A}}}\,\widetilde{\partial})\,(\partial A \;-\; \widetilde{A}\,\widetilde{\partial}) \;+\; \underline{\underline{\partial}}_{\widetilde{A}}\,(\partial A \;-\; \widetilde{A}\,\widetilde{\partial})\,(\partial\underline{\underline{A}} \;-\; \underline{\underline{\widetilde{A}}}\,\widetilde{\partial})\,. \qquad (3.53)$$

In the latter expression, we have underlined the particular instances of field $A$ which are acted upon by $\partial_{\widetilde{A}}$ with a double line. All other occurrences of $A$ are considered constant for the matter of variation. The space-time derivatives $\partial$ here act on the factors of $A$ that are closest to them. We need to free the underlined variables $A$ from the space-time derivatives, by integrating the latter by parts. This will change the sign in front of each of these derivatives, but will not move them anywhere because they are quaternionic. To indicate that they are now acting on different targets, we will underline them, as well as their targets, using a single line,

$$\underline{\underline{\partial}}_{\widetilde{A}}\,\big(-\,\partial\underline{A} \;+\; \underline{\widetilde{A}}\,\partial\big)\big(\underline{\hspace{1.2cm}\cdots\hspace{1.2cm}}\big) \;+\; \underline{\underline{\partial}}_{\widetilde{A}}\,\big(\underline{\hspace{1.2cm}\cdots\hspace{1.2cm}}\big)\big(-\,\partial\underline{A} \;+\; \underline{\widetilde{A}}\,\partial\big)\,. \qquad (3.54) \quad \texttt{\{by-parts\}}$$

19

Here the dots symbolize the term

$$\partial\, A \;-\; \widetilde{A}\, \widetilde{\partial}\,. \tag{3.55}$$ `{dots-term}`

For the sake of variation, the space-time derivatives are now just constants. In fact, everything is constant in Eq. (3.54) in regards to varying with respect to $A$ — except for the doubly-underlined $A$ factors. Applying identities (B.45) and (B.46), we get, explicitly

$$2\,\widetilde{\partial}\,(\quad\ldots\quad) \;+\; 4\,\widetilde{\partial}\,(\quad\ldots\quad) \;+\; 2\,\widetilde{\partial}\,(\quad\ldots\quad)^{\widetilde{\;}} \;+\; 4\,(\quad\ldots\quad)_0\,\widetilde{\partial}\,. \tag{3.56}$$ `{varied}`

The third term here involves the quaternionic conjugate of (3.55), while the last term involves the real component of the (3.55). Since (3.55) is purely imaginary, applying quaternionic conjugation just changes its sign, while taking the real part annihilates it,

$$4\,\widetilde{\partial}\,\big(\partial\, A \;-\; \widetilde{A}\,\widetilde{\partial}\big) \;=\; -\,8\,\widetilde{\partial}\,\Phi\,. \tag{3.57}$$

Although a fair exercise in differentiation, the derivative of the conjugate term $\overline{\Phi}{}^2$ will actually give the same contribution. The reason we can guess that is that the two terms $\Phi^2$ and $\overline{\Phi}{}^2$ only differ by their complex-imaginary part

$$\mathrm{Im}\,\Phi^2 \;\;\propto\;\; \vec{E}\circ\vec{B}\,, \tag{3.58}$$

which is a boundary term and thus does not affect the equations of motion.

Altogether, we arrive to,

$$\partial_{\widetilde{A}}\,\mathcal{L}_{\mathrm{Maxwell}} \;=\; j \;-\; \widetilde{\partial}\,\Phi\,, \tag{3.59}$$

or,

$$\widetilde{\partial}\,\Phi \;=\; j\,, \tag{3.60}$$ `{Maxwell}`

which is the *quaternionic Maxwell's equation*. Notice that, because it is a complex-quaternionic equation, it includes all eight real Maxwell's equations. The reader is encouraged to check that they are reproduced correctly in Eq. (3.60).

# 4   Conclusions

We have demonstrated the construction of Dirac and Weyl spinors in the complex quaternionic space $\mathcal{C}\otimes\mathcal{H}$. The spinors become part of the algebra, on the same

<div align="center">20</div>

grounds as the kinetic operators acting on them. This allows for finding purely-algebraic solutions of the Dirac equation in various settings, without resorting to the matrix form. Different types of conjugation of spinors (2.5), (2.6) and (2.7) is what particularly distinguishes the quaternionic formalism from the regular matrix description. For example, the discrete transformations (A.19)–(A.22) have a simpler appearance. In addition we notice that these complex and quaternionic conjugations are not as easy to implement in the regular spinor formalism.

After building the Lagrangian, we have been able to develop a scheme of using the variational calculus to consistently derive the equations of motion. This calculus is well applicable to spinors and to vectors, which we have demonstrated by deriving the (single) Maxwell's equation.

An important result is the digression of the algebra in the non-relativistic limit. By taking this limit, we have established the form of the Schrödinger's equation in the quaternionic formulation. Quantum Mechanics in the quaternionic form has been a subject of long study with various success. We show that it is possible to uniquely fix this form by starting from the Dirac equation. No extra degrees off freedom appear. Furthermore, the existence of the spin of the wavefunction is the natural consequence of the latter being quaternionic. The solution can be sought in the *algebraic* form. Then, by applying the correct projector one finds the quantum-mechanical wavefunction. Such solutions will be the subject of further study. Once the spin is discarded, the Schrödinger's equation takes the usual *complex* form. The quaternionic algebra reduces to the complex one in the non-relativistic limit. Extending this hypothesis, we hope to expect that addition of the strong and weak interactions grows the algebra to a unifying $\mathcal{O} \otimes \mathcal{H} \otimes \mathcal{C} \otimes \mathcal{R}$.

In this context, we can address the length of the Lagrangian of electrodynamics (2.22). As we have mentioned, it is possible to write it in a more compact way. There have been multiple reasons we have not done this in this paper. Besides the desired non-relativistic limit, we note that the Standard Model itself treats the chiralities differently. Finally, we believe, that the right answer to a compact form is given by placing the theory into the framework of electroweak interactions. This comprises a promising direction for future work.

# Acknowledgements

# A  Spinors

So how do we deduce the bases (1.27) and (1.32) for spinors? In this section we essentially expand on the construction introduced in [*thesis*]. We start from the correspondence

$$\sigma_1, \quad \sigma_2, \quad \sigma_3 \quad \rightarrow \quad i\,\hat{\imath}, \quad i\,\hat{\jmath}, \quad i\,\hat{k}, \tag{A.1}$$

for the Pauli matrices. This is a very natural association, and the only other reasonably alternative choice here would be a different sign on the right-hand side.

Next we build the spin-up and spin-down states, assuming the rest frame of reference. The operator of the canonical $z$-component of spin should be

$$\frac{\sigma^3}{2} \quad \rightarrow \quad \frac{i\,\hat{k}}{2},$$

so the up- and down-states should satisfy

$$i\,\hat{k}\,\psi_\uparrow \quad = \quad +\,\psi_\uparrow \tag{A.2}$$

and

$$i\,\hat{k}\,\psi_\downarrow \quad = \quad -\,\psi_\downarrow. \tag{A.3}$$

Or, in other words,

$$(1 \mp i\,\hat{k})\,\psi_{\uparrow,\downarrow} \quad = \quad 0. \tag{A.4}\quad\text{\{updown\}}$$

The reason that a product of two complex quaternions can vanish, is because the algebra of complex quaternions does not admit a positive-definite norm. We can still use the usual quaternionic norm,

$$\|\,a\,\|^2 \quad = \quad a_0^2 \quad + \quad a_1^2 \quad + \quad a_2^2 \quad + \quad a_3^2, \tag{A.5}\quad\text{\{qnorm\}}$$

ignoring the fact that $a_0$, ... are complex numbers. Such a norm will be multiplicative, but not positive-definite. The reason the product of the two factors in Eq. (A.4) vanishes is because the norm of at least one of those factors vanishes. Indeed, the norm of $1 \mp i\,\hat{k}$ as calculated via Eq. (A.5) is zero.

So how about $\psi_{\uparrow,\downarrow}$? How many solutions can there be? These questions are easily answered by noticing that objects $1 + i\,\hat{k}$ and $1 - i\,\hat{k}$ are, in fact, projectors:

$$P_L \quad = \quad \frac{1 + i\,\hat{k}}{2}, \qquad\qquad P_L^2 \quad = \quad P_L,$$

$$P_R \quad = \quad \frac{1 - i\,\hat{k}}{2}, \qquad\qquad P_R^2 \quad = \quad P_R, \tag{A.6}$$

23

which we judiciously have named the left- and right-handed projectors. Indeed, by multiplying a quaternion on the left (or equally well, on the right) by $P_L$, we are obviously performing a linear operation upon the components of that quaternion. The square of such an operation equals the operation itself $\Rightarrow$ the operation is a projection. Obviously the same is true for $P_R$.

That means that if we run $\psi$ through all values of complex quaternions, the product $P_L\psi$ will span only a portion of the quaternionic space — an ideal. What fraction of the entire algebra does it span? We notice that $P_L$ and $P_R$ are complementary projectors:

$$ P_R \;=\; P_L^* , \qquad\qquad P_L \;+\; P_R \;=\; 1 . \qquad (A.7) $$

Because they are symmetric, they can only project equal-size subsets of the algebra. And since they are complementary to each other, the union of those subsets must comprise the entire algebra. In other words, $P_L$ and $P_R$ split the algebra in two halves.

Now, equation (A.4) can re-written as,

$$ P_{R,L}\,\psi_{\uparrow,\downarrow} \;=\; 0 , \qquad\qquad\qquad (A.8) \quad \{\text{PRL}\} $$

meaning that the most general $\psi_\uparrow$ must sit in one half of the algebra, while the most general $\psi_\downarrow$ must sit in the other half. Since a generic complex quaternion has four complex components, there are two complex solutions for $\psi_\uparrow$ and as many for $\psi_\downarrow$. We have to stress here that these halves are *not identified* with the left- and right-handed chiral spaces. For this reason we have not given names to these subspaces, other than "spin-up" and "spin-down" spaces.

The easiest solution to (A.8) is given by the orthogonality of the projectors:

$$ P_R\,P_L \;=\; 0 , \qquad\qquad P_L\,P_R \;=\; 0 . \qquad (A.9) $$

So, seemingly $P_L$ could be identified with a spin-up state, and $P_R$ — with the corresponding spin-down state. However, these are states of different chiralities. Indeed, since $P_L \;=\; P_R^*$, such "spinors" transform in the mutually-conjugate representations of the Lorentz group. This fact puts them into the opposite chirality spaces.

Let us summarize our goal and achievements now. We are looking for four complex states $\psi_{L\uparrow}$, $\psi_{L\downarrow}$, $\psi_{R\uparrow}$ and $\psi_{R\downarrow}$. We have already found

$$ \psi_{L\uparrow} \;=\; P_L , \qquad\qquad \psi_{R\downarrow} \;=\; P_R . \qquad (A.10) $$

24

It is not difficult to find the other two states. For example, if one wishes to avoid pure guessing, which would perfectly work here too, we know that matrix $-i\sigma^2$ turns a spin-up state into a spin-down state:

$$-i\,\sigma^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This exactly corresponds to multiplying by $\hat{\jmath}$ on the left, and we arrive to

$$\psi_{L\downarrow} = \hat{\jmath}\,P_L\,, \qquad\qquad \psi_{R\uparrow} = -\hat{\jmath}\,P_R\,. \qquad\qquad \text{(A.11)}$$

So we recap that the *left-handed* spinors can be written as

$$\psi_L = \xi_L\,P_L + \chi_L\,\hat{\jmath}\,P_L\,, \qquad\qquad \text{(A.12)} \quad \texttt{\{lbasis-again\}}$$

while the *right-handed* spinors are represented as

$$\psi_R = -\xi_R\,\hat{\jmath}\,P_R + \chi_R\,P_R\,. \qquad\qquad \text{(A.13)} \quad \texttt{\{rbasis-again\}}$$

A few important comments are in order here. The chiral subspaces are defined here by multiplying arbitrary quaternions by projectors $P_L$ or $P_R$ on the *right*. In other words, $\psi\,P_L$ spans the set of all left-handed spinors, and $\psi\,P_R$ of all right-handed. This way, right multiplication splits the set of quaternions into the two chirality subspaces. Whereas, *left* multiplication splits the quaternions into spin-up and spin-down subspaces, because *e.g.*

$$P_R\,(P_L\,\psi) = 0\,.$$

It is a trivial fact now that a set of the type $\{\,P_L\,\psi\,\}$ or $\{\,\psi\,P_L\,\}$ forms a left (right) ideal, since for an arbitrary quaternion $a$, the product, say,

$$a\,(\psi\,P_L) = (a\,\psi)\,P_L\,,$$

resides in the same subspace as $\psi\,P_L$.

The other important remark is about the amount of freedom that we have in defining the bases via Eqs. (A.12), (A.13). We have freedom in choosing the component of the spin to be measurable — for which we chose $\hat{k}$. The other freedom was in parametrizing the spin-down component of $\psi_L$, for which we chose $\hat{\jmath}$ as an

25

orthogonal direction. Overall, we could have chosen any two unit vectors $\hat{a}$ and $\hat{b}$ in place of $\hat{k}$ and $\hat{j}$, subject to the only restriction

$$\hat{a} \cdot \hat{b} \;=\; 0 \,.$$

The *standard representation* for the spinors can be established as follows. We notice that multiplying a left-handed spinor $\psi_L$ on the right by $\hat{j}$ brings it into the right-handed subspace,

$$\psi_L \, \hat{j} \;\sim\; -\,\psi_R \,, \tag{A.14}$$

and, respectively, a right-handed spinor $\psi_R$ into the left-handed space,

$$\psi_R \, \hat{j} \;\sim\; \psi_L \,. \tag{A.15}$$

This way, by forming a linear combination

$$\frac{\psi_D \;+\; \psi_D \, \hat{j}}{\sqrt{2}} \;=\; \psi_D \, \frac{1 \;+\; \hat{j}}{\sqrt{2}} \,, \tag{A.16}$$

we create a spinor consisting of

$$\zeta \;=\; \frac{\psi_L + \psi_R}{\sqrt{2}} \qquad \text{and} \qquad -\,\eta \;=\; \frac{\psi_R - \psi_L}{\sqrt{2}} \,. \tag{A.17}$$

Here the sum and difference are understood in terms of the usual spinor components. In terms of the quaternionic addition, spinors $\zeta$ and $\eta$ can be written *e.g.* as

$$\zeta \;=\; \frac{\psi_L + \psi_R \hat{j}}{\sqrt{2}} \,, \qquad\qquad \eta \;=\; \frac{\psi_L - \psi_R \hat{j}}{\sqrt{2}} \,. \tag{A.18}$$

In this example they both reside in the left-handed subspace — for definiteness, but this has no special rôle in the standard representation, and the corresponding right-handed expressions can be readily written.

As was discussed in [*thesis*], the Dirac algebra $\mathrm{C}\ell(4) \sim \mathrm{C}\ell(2) \otimes \mathrm{C}\ell(2)$ is realized in a very interesting way on complex quaternions. One factor of $\mathrm{C}\ell(2)$ acts on $\psi_D$ via multiplication on the *left* while the other one — via multiplication on the *right*. Left multiplication by quaternions rotates the spin components in $\psi_L$ and $\psi_R$ independently, while right multiplication does not rotate the spin components, and instead rotates $\psi_L \leftrightarrow \psi_R$. This latter rotation can be used to identify the discrete symmetries.

26

## A.1 Discrete symmetries

Here we will discuss the action of $C$, $P$ and $T$ symmetries on fermions. These symmetries can be conveniently written for an entire Dirac fermion $\psi_D$. We use here the phase conventions of $[BLP]$.

· Charge conjugation $C$ is realized by complex conjugation of the fermion,

$$\psi_D(x) \;\to\; i\,\psi_D^*(x)\,. \tag{A.19}$$

· Parity transformation $P$ is given by right multiplication by $-\hat{\imath}$,

$$\psi_D(t,\,\vec{x}) \;\to\; -\,\psi_D(t,\,-\vec{x})\,\hat{\imath}\,, \tag{A.20}$$

which interchanges the chiral components as,

$$\psi_L \;\to\; i\,\psi_R\,, \qquad\qquad \psi_R \;\to\; i\,\psi_L\,. \tag{A.21}$$

· Time inversion $T$ is performed by complex conjugation and multiplying by $i\hat{\jmath}$ on the right,

$$\psi_D(t,\,\vec{x}) \;\to\; i\,\psi_D^*(-t,\,\vec{x})\,\hat{\jmath}\,. \tag{A.22}$$

All three transformations result in

$$C\,P\,T\,\psi_D(x) \;=\; \psi_D(-x)\,\hat{k}\,. \tag{A.23}$$

Right-multiplying by $i\hat{k}$ preserves the sign of the left-handed spinor, while flips the sign of the right-handed part. Therefore, multiplying by $\hat{k} = (-i)i\hat{k}$ corresponds to multiplying by $i\gamma^5$ in the spinor representation. Notice that the above transformations have a significantly simpler form than they do in the $\gamma$-matrix representation.

# B  Differentiation

Derivation of the equations of motion from the Lagrangian involves one crucial step — variation, which we will loosely call differentiation. Indeed, the problem of extremizing the action essentially reduces to differentiating with respect to a quaternion. As soon as we are able to differentiate, variational calculus sets in place. We start with regular quaternions.

27

## B.1 Ordinary quaternions

It is well known that there is no existing analyticity theory of quaternions analogous to that of complex numbers. Even the notion of a derivative is not well established. While complex analyticity rests on holomorphic functions that depend on $z$ and are independent of $\overline{z}$,

$$f(z, \overline{z}) \;=\; f(z),$$

this property cannot be extended to quaternions.

There is one crucial reason for this: if a complex function depends on $z$, there is no way that it can be represented as a function of $\overline{z}$. In other words, variable $z$ cannot be converted into $\overline{z}$ by multiplying it by any constants. This is *not so* with quaternions. For, given a quaternion $q$, we find that a combination

$$-\frac{q + \hat{\imath}\,q\,\hat{\imath} + \hat{\jmath}\,q\,\hat{\jmath} + \hat{k}\,q\,\hat{k}}{2} \;=\; \widetilde{q} \tag{B.24}$$

is exactly the conjugate of $q$. Any other type of conjugation that can be introduced for quaternions (say, the one that only flips the sign of a single imaginary unit $\hat{\imath}$) can also be represented in such an "arithmetic" form. That means that any function $f(q)$ can be viewed as a function of $\widetilde{q}$, just with different coefficients. So the notion of holomorphy cannot be applied to quaternions, at least directly.

It is, however, possible to define a convenient notion of a derivative. It works especially well if the function being differentiated is real (although it does not have to be).

Let us say that $f(q)$ is a function of variable $q$,

$$q \;\equiv\; t \;+\; \vec{r} \;=\; t \;+\; x\,\hat{\imath} \;+\; y\,\hat{\jmath} \;+\; z\,\hat{k}. \tag{B.25}$$

We introduce a derivative $\partial$,

$$\partial \;\equiv\; \partial_t \;+\; \vec{\nabla} \;=\; \partial_t \;+\; \hat{\imath}\,\partial_x \;+\; \hat{\jmath}\,\partial_y \;+\; \hat{k}\,\partial_z. \tag{B.26} \quad \text{\{diff-d\}}$$

Strictly speaking, in analogy to complex numbers, we should call this quantity a conjugate derivative $2\,\widetilde{\partial}$ multiplied by a factor of two or so, as we will see below. However, to keep notations flat and straightforward, we just call it $\partial$, and we define a conjugate derivative as

$$\widetilde{\partial} \;\equiv\; \partial_t \;-\; \vec{\nabla} \;=\; \partial_t \;-\; \hat{\imath}\,\partial_x \;-\; \hat{\jmath}\,\partial_y \;-\; \hat{k}\,\partial_z. \tag{B.27} \quad \text{\{diff-d-conj\}}$$

Let $a$ denote an arbitrary quaternionic constant. Then the following key identities

$$\partial q = -2 \qquad\qquad \partial \widetilde{q} = 4$$

$$\widetilde{\partial} q = 4 \qquad\qquad \widetilde{\partial}\widetilde{q} = -2, \qquad\qquad \text{(B.28)} \quad \{\texttt{dq}\}$$

and

$$\partial (aq) = -2\,\widetilde{a} \qquad\qquad \partial (a\widetilde{q}) = 4\,a_0$$

$$\widetilde{\partial}(aq) = 4\,a_0 \qquad\qquad \widetilde{\partial}(a\widetilde{q}) = -2\,\widetilde{a}. \qquad\qquad \text{(B.29)} \quad \{\texttt{daq}\}$$

enable us to perform differentiation. Note that the set of identities (B.28) applies equally well when the derivative is acting from the right.

Any function of $q$ which can be represented as a power series (in fact, any *analytical* function of coordinates $t$, $x$, $y$, $z$),

$$f(q) = \ldots + \alpha\, q\, \beta\, q\, \gamma\, q\, \delta + \ldots \qquad\qquad \text{(B.30)} \quad \{\texttt{q-series}\}$$

can now be differentiated using Eqs. (B.28), (B.29). It is sufficient to define the action of the derivative[2] on the monomial in (B.30). Derivative $\partial$ will act on each factor of $q$ in (B.30) in turn. Let us underline each factor when it is differentiated, and when it is not — it can be considered a constant:

$$\partial\,\alpha\, q\, \beta\, q\, \gamma\, q\, \delta = \underline{\partial}\,\alpha\,\underline{q}\,\beta\, q\,\gamma\, q\,\delta + \underline{\partial}\,\alpha\, q\,\beta\,\underline{q}\,\gamma\, q\,\delta + \underline{\partial}\,\alpha\, q\,\beta\, q\,\gamma\,\underline{q}\,\delta. \qquad \text{(B.31)} \quad \{\texttt{diff-term}\}$$

We are also underlining the derivative operator, so that in the case when there is more than one derivative is present, it is clear which one acts on what variable. Every factor in (B.31) that is not differentiated can be considered a constant. Thus, we can use the first one of identities (B.29),

$$\partial\,\alpha\, q\,\beta\, q\,\gamma\, q\,\delta = -2\,\widetilde{\alpha}\,\beta\, q\,\gamma\, q\,\delta - 2\,\widetilde{\beta}\,\widetilde{q}\,\widetilde{\alpha}\,\gamma\, q\,\delta - 2\,\widetilde{\gamma}\,\widetilde{q}\,\widetilde{\beta}\,\widetilde{q}\,\widetilde{\alpha}\,\delta.$$

The efficiency of this approach is that we did not have to deal with components.

In which sense is $\partial f(q)$ a derivative of $f(q)$? Can we restore an infinitesimal change $\Delta f$ due the increment $\Delta q$, using this derivative? The answer is not quite straightforward, as the correct expression is given by

$$\Delta f = \underline{f}(q)\,\frac{\widetilde{\partial}\,\Delta q + \Delta\widetilde{q}\,\partial}{2}. \qquad\qquad \text{(B.32)} \quad \{\texttt{delta-f}\}$$

---

[2]We are not really "defining" this action, as it is explicitly defined in Eq. (B.26), but instead giving a recipe how to efficiently compute such a derivative.

29

This result is obvious because the fraction here actually equals $\partial_\mu \Delta q^\mu$. Being a scalar, it can be written on either side of $f(q)$. The problem here is that one of the two terms in that fraction will necessarily have a derivative operator $\partial$ sitting furthest away from $f(q)$,

$$f(q)\,\Delta\widetilde{q}\,\partial\,,$$

and so will not form a derivative in our sense. In other words, here the derivative has to depend on the increment $\Delta q$ in order to correctly reproduce $\Delta f$.

Instead of engaging with this problem, let us switch to *real* functions $f(q)$ — our prime subject of interest — since the Lagrangian is real. For real $f(q)$, the rightmost term in the fraction in (B.32) commutes with $f(q)$, and so

$$\Delta f \;=\; \frac{1}{2}\left(\Delta\widetilde{q}\,\partial\,f(q)\;+\;f(q)\,\widetilde{\partial}\,\Delta q\right),$$

where we also have replaced $f(q)$ with $\widetilde{f}(q)$ in the first term in the bracket. The extremum of $f(q)$ is obviously achieved when

$$\partial\,f(q) \;=\; 0\,, \tag{B.33}$$

since this also implies that $\widetilde{f}(q)\,\widetilde{\partial} \;=\; 0$.

Let us see how this works in practice. Let us apply a derivative operator $\partial$ on a real function $f(q)$. Even if we do not know the form of $f(q)$, when viewed as a series it can always be represented as

$$f(q) \;=\; \ldots \;+\; q\,a \;+\; \widetilde{a}\,\widetilde{q} \;+\; \ldots\,, \tag{B.34} \quad \text{\{f-real\}}$$

since when we differentiate a particular factor of $q$ in a given monomial term, all the other factors of $q$ in that term can be considered constant and absorbed into $a$. Applying identities (B.28), (B.29) to function (B.34), we get

$$\partial\,f(q) \;\supset\; -\,2\,a \;+\; 4\,a_0 \;=\; 2\,\widetilde{a}\,. \tag{B.35} \quad \text{\{der-f\}}$$

Note that we would have arrived to the same result had we written $q$ and $a$ in Eq. (B.34) in a different order,

$$f(q) \;=\; \ldots \;+\; a\,q \;+\; \widetilde{q}\,\widetilde{a} \;+\; \ldots\,, \tag{B.36}$$

only now the other identities in (B.28), (B.29) would have been involved.

Now let us apply operator $\widetilde{\partial}$ to Eq. (B.34). We find,

$$\widetilde{\partial} f(q) \;=\; \ldots \;+\; 4\,a \;-\; 2\,a \;+\; \ldots \;=\; \ldots \;+\; 2\,a \;+\; \ldots. \qquad \text{(B.37)} \quad \{\texttt{der-f-conj}\}$$

Equations (B.35) and (B.37) tell us that $(1/2)\,\partial$ in fact differentiates with respect $\widetilde{q}$, while $(1/2)\,\widetilde{\partial}$ — with respect to $q$. This is exactly the same as with complex numbers, even with a matching factor of $1/2$. Still, we would like to leave our definitions of $\partial$ and $\widetilde{\partial}$ as they are in order to keep notations plain.

Being able to find the extremum of a real function, we now know how to derive the extremum of the action — that is, the equations of motion. The derivative gets promoted to a variational derivative.

## B.2 Generic complex quaternions

Does the above picture of differentiation change for complex quaternions? If we are talking about generic complex quaternions — not really. Generic complex quaternions have eight real degrees of freedom. Everything stays the same, taking into account that the components $t$, $x$, $y$ and $z$ are now *complex numbers*.

The derivatives

$$\partial \;=\; \partial_t \;+\; \hat{\imath}\,\partial_x \;+\; \hat{\jmath}\,\partial_y \;+\; \hat{k}\,\partial_z\,,$$

$$\widetilde{\partial} \;=\; \partial_t \;-\; \hat{\imath}\,\partial_x \;-\; \hat{\jmath}\,\partial_y \;-\; \hat{k}\,\partial_z \qquad \text{(B.38)} \quad \{\texttt{der-cquat}\}$$

stay exactly what they are. Here, of course, we assume that the complex derivatives are defined in the usual way,

$$\partial_t \;=\; \frac{1}{2}\left\{\,\partial_{t_1} \;-\; i\,\partial_{t_2}\,\right\}, \qquad \ldots. \qquad \text{(B.39)}$$

We also have the two conjugates of (B.38),

$$\partial^* \;=\; \partial_t^* \;+\; \hat{\imath}\,\partial_x^* \;+\; \hat{\jmath}\,\partial_y^* \;+\; \hat{k}\,\partial_z^*\,,$$

$$\partial^\dagger \;=\; \partial_t^* \;-\; \hat{\imath}\,\partial_x^* \;-\; \hat{\jmath}\,\partial_y^* \;-\; \hat{k}\,\partial_z^*. \qquad \text{(B.40)} \quad \{\texttt{der-cquat-con}$$

Obviously, any function of $q$ (and not $q^*$) will have derivatives $\partial^*$ and $\partial^\dagger$ vanishing on it.

Relations (B.28) and (B.29) are valid in our complex case verbatim. There is also a complex conjugated copy of them. We are not going to pay much attention to these relations.

We are able to re-use literally all the formulas applicable to the differentiation with respect to regular quaternions, because complex quaternions have all eight degrees of freedom occupied.

The story becomes more interesting when we have to address *constrained* quaternions — *e.g.* vectors or spinors. We call them constrained, because not all eight components in them are independent (or non-zero). Thus, the derivatives (B.38), (B.40) cannot be applied to them verbatim to produce a meaningful result.

There is also the field strength $\Phi$ which is a constrained quaternion, but we normally do not vary with respect field strengths.

## B.3   Constrained complex quaternions — vectors

Vectors open up quite an interesting story, as the only physical vector field variables are gauge fields. Thus, differentiation with respect to vectors enables us to derive an expression for the current, as well as to formally derive the form of the quaternionic Maxwell's equation.

Formally vectors can be defined by the constraint condition

$$q^{\dagger} \;=\; q \,, \tag{B.41}$$

which means that only four real components of it are non-zero:

$$q \;\equiv\; t \;+\; i\vec{r} \;=\; t \;+\; i\left(\hat{i}\,x \;+\; \hat{j}\,y \;+\; \hat{k}\,z\right). \tag{B.42}$$

In relation to it we define the derivative operators

$$\partial \;\equiv\; \partial_t \;+\; i\vec{\nabla} \;=\; \partial_t \;+\; i\left(\hat{i}\,\partial_x \;+\; \hat{i}\,\partial_y \;+\; \hat{k}\,\partial_z\right) \tag{B.43}$$

and

$$\widetilde{\partial} \;=\; \partial^{*} \;=\; \partial_t \;-\; i\left(\hat{i}\,\partial_x \;+\; \hat{i}\,\partial_y \;+\; \hat{k}\,\partial_z\right). \tag{B.44}$$

Relations (B.28) and (B.29) are valid but now their right-hand sides are interchanged vertically,

$$\partial\,q \;=\; 4 \qquad\qquad \partial\,\widetilde{q} \;=\; -2$$

$$\widetilde{\partial}\,q \;=\; -2 \qquad\qquad \widetilde{\partial}\,\widetilde{q} \;=\; 4\,, \tag{B.45} \quad \{\texttt{dq-vector}\}$$

and

$$\partial\,(aq) \;=\; 4\,a_0 \qquad\qquad \partial\,(a\widetilde{q}) \;=\; -\,2\,\widetilde{a}$$

$$\widetilde{\partial}\,(aq) \;=\; -\,2\,\widetilde{a} \qquad\qquad \widetilde{\partial}\,(a\widetilde{q}) \;=\; 4\,a_0\,. \qquad\qquad \text{(B.46)} \quad \text{\{daq-vector\}}$$

Other than that, there is no difference from differentiating with respect to regular quaternions, and these formulas are enough to derive the equations of motion.

## B.4   Constrained complex quaternions — spinors

Full Dirac spinors are just complex quaternions, and so differentiation with respect to them is in fact addressed by operators (B.38). There is nothing new from that perspective. One only has to keep in mind that spinors are Grassmann numbers.

Weyl spinors, however are constrained. Even though a single Weyl spinor $\psi_L$ or $\psi_R$ has all eight components non-zero, it only has four independent ones.

Let us take a left-handed fermion $\psi_L$,

$$\psi_L \;=\; \xi_L\,P_L \;+\; \chi_L\,\hat{\jmath}\,P_L\,. \qquad\qquad \text{(B.47)}$$

It turns out, that differentiation with respect to $\psi_L$ is only meaningful when done from the right. The reason behind is that when multiplying by anything from the left, it is not possible to "cancel" the projectors $P_L$. We have,

$$\underline{\psi_L}\left(\partial_{\xi_L} \,-\, \hat{\jmath}\,\partial_{\chi_L}\right) \;=\; P_L \;+\; P_R \;=\; 1\,. \qquad\qquad \text{(B.48)} \quad \text{\{diff-psiL\}}$$

Again, we have underlined $\psi_L$ to show that the derivatives act to the left. Here $\partial_{\xi_L}$ and $\partial_{\chi_L}$ are the usual Grassmann derivatives with respect to the complex components $\xi_L$ and $\chi_L$ correspondingly. Analogously we find

$$\underline{\psi_R}\left(\hat{\jmath}\,\partial_{\xi_R} \,+\, \partial_{\chi_R}\right) \;=\; \underline{\psi_R}\,\hat{\jmath}\left(\partial_{\xi_R} \,-\, \hat{\jmath}\,\partial_{\chi_R}\right) \;=\; 1\,. \qquad\qquad \text{(B.49)} \quad \text{\{diff-psiR\}}$$

So we essentially have identified the spinor derivatives, we just need to re-write them in terms of the quaternionic components $t$, $x$, $y$ and $z$. Although we are really considering $\psi_L$ and $\psi_R$ separately, let us combine them both into a single Dirac fermion. This is only done for the sake of completeness — a derivative with respect

33

to $\psi_L$ does not touch $\psi_R$ in any sense, because the components of the two spinors are independent complex numbers. We have,

$$\psi_D = \frac{\xi_L + \chi_R}{2} + \hat{\imath}\, i\, \frac{\xi_R + \chi_L}{2} + \hat{\jmath}\, \frac{-\xi_R + \chi_L}{2} + \hat{k}\, i\, \frac{\xi_L - \chi_R}{2}$$

$$= t + \hat{\imath}\, x + \hat{\jmath}\, y + \hat{k}\, z. \qquad (B.50) \quad \{\texttt{psiD-as-quat}\}$$

Knowing this, we can express derivatives $\partial_{\xi_L}$, *etc* in terms of the derivatives with respect to components $t$, $x$, $y$ and $z$. Plugging them into the expressions in parentheses in equations (B.48) and (B.49) we find,

$$\partial_{\psi_L} \propto \frac{\partial_t - i\hat{\jmath}\,\partial_x - \hat{\jmath}\,\partial_y + i\,\partial_z}{2},$$

$$\partial_{\psi_R} \propto \frac{\partial_t + i\hat{\jmath}\,\partial_x - \hat{\jmath}\,\partial_y - i\,\partial_z}{2}. \qquad (B.51) \quad \{\texttt{prelim-psi-de}$$

Although we could write equality signs in these equations, we postpone this until we do the last extra step. Namely, equations (B.51) do not seem to be very aesthetic, or easy to remember. Notice, that the action of these derivatives will not change if we multiply them by the projectors $P_L$ and $P_R$ on the left, correspondingly. Indeed, these projectors will carve out $\psi_L$ or $\psi_R$ from Eq. (B.50), before the actual derivatives will hit them. It is then a matter of a simple multiplication to observe that the right-hand sides of Eq. (B.51) when multiplied by the corresponding projectors, match a regular derivative $\widetilde{\partial}$ multiplied by the same projectors

$$\partial_{\psi_L} = \frac{1}{2}\, P_L\, \widetilde{\partial},$$

$$\partial_{\psi_R} = \frac{1}{2}\, P_R\, \widetilde{\partial}. \qquad (B.52) \quad \{\texttt{spinor-der}\}$$

where $\widetilde{\partial}$ is the derivative (B.38) with respect to the components (B.50) of $\psi_D$ viewed as a complex quaternion,

$$\widetilde{\partial} = \partial_t - \hat{\imath}\,\partial_x - \hat{\jmath}\,\partial_y - \hat{k}\,\partial_z. \qquad (B.53)$$

From Eq. (B.52),

$$\partial_{\psi_L} + \partial_{\psi_R} = \frac{1}{2}\,\widetilde{\partial}. \qquad (B.54) \quad \{\texttt{spinor-der-to}$$

34

We remind that all derivatives in Eqs. (B.52) and (B.54) act to the *left*.

Quite analogously, we consider the hermitean-conjugate spinors $\psi_L^\dagger$ and $\psi_R^\dagger$, for which

$$\partial_{\psi_L^\dagger} \quad \propto \quad \partial_{\xi_L^*} \quad + \quad \hat{j}\,\partial_{\chi_L^*}\,,$$

$$\partial_{\psi_R^\dagger} \quad \propto \quad \left( \partial_{\xi_R^*} \quad + \quad \hat{j}\,\partial_{\chi_R^*} \right)\,(-\hat{j})\,, \tag{B.55} \quad \texttt{\{diff-psiLR-da}$$

where the derivatives now act *to the right*. Notice that the expressions in (B.55) are precisely the hermitean conjugates of those in Eqs. (B.48) and (B.49). Proceeding the same way as we arrived to Eq. (B.51) and past it, we, now unsurprisingly, find

$$\partial_{\psi_L^\dagger} \quad = \quad \frac{1}{2}\,\partial^*\,P_L\,,$$

$$\partial_{\psi_R^\dagger} \quad = \quad \frac{1}{2}\,\partial^*\,P_R\,. \tag{B.56} \quad \texttt{\{spinor-der-da}$$

We stress again that here the derivatives act to the *right*. Operator $\partial^*$ is the derivative with respect to the complex conjugates of the components (B.50) of $\psi_D$, the latter viewed as a complex quaternion,

$$\partial^* \quad = \quad \partial_{t^*} \quad - \quad \hat{i}\,\partial_{x^*} \quad - \quad \hat{j}\,\partial_{y^*} \quad - \quad \hat{k}\,\partial_{z^*}\,,$$

$$\psi_D^* \quad = \quad t^* \quad + \quad \hat{i}\,x^* \quad + \quad \hat{j}\,y^* \quad + \quad \hat{k}\,z^*\,. \tag{B.57}$$

Equations (B.52) and (B.56) is our final answer for the spinor derivatives. The reason is that we already now how to apply a differential operator $\partial$ or its conjugates in an efficient way — we have equations (B.28) and (B.29) at our disposal.

If an expression is linear, say in $\psi_L$, then differentiation simply "erases" the latter from the expression (since the Lagrangian is real, one can always arrange $\psi_L$ to stand on the right of the expression). Fermionic expressions are almost always linear in spinors ($\psi_L^\dagger$ or $\psi_L^*$ is an independent variable from $\psi_L$), except for four-fermion interactions. Even then, Appendix B.1 explains how to deal with such functions.

As was shown in Appendix B.1, an extremum of a real function is attained when

$$\partial f \quad = \quad 0\,. \tag{B.58} \quad \texttt{\{extremum-eq\}}$$

Whether the quaternions are ordinary or complex, does not change this. As we saw above, such a derivative can be split into left- and right-handed parts, each of which

provides an independent equation — for $\psi_L$ and $\psi_R$ correspondingly. To see that, it is enough to multiply Eq. (B.58) on the right by a left- or right-handed projector. Thus, we have a recipe for deriving the equations of motion for Weyl spinors as well as for Dirac spinors.

Let us give simple example that we use — a Lagrangian of a Dirac fermion,

$$\mathcal{L} \;=\; \psi_L^\dagger \, i\mathcal{D}\,\psi_L \;+\; \psi_R^\dagger \, i\widetilde{\mathcal{D}}\,\psi_R \;-\; \left( m\,\psi_L^\dagger \, \psi_R \, \hat{\jmath} \;+\; \text{q.c.} \right) \;+\; \text{c.c.} \qquad \text{(B.59)}$$

We will need this Lagrangian in the full form,

$$\mathcal{L} \;=\; \psi_L^\dagger \, i\mathcal{D}\,\psi_L \;+\; \psi_R^\dagger \, i\widetilde{\mathcal{D}}\,\psi_R \;+\; \widetilde{\psi}_L \, i\mathcal{D}^*\,\psi_L^* \;+\; \widetilde{\psi}_R \, i\mathcal{D}^\dagger\,\psi_R^* \;-$$
$$-\; m\left( \psi_L^\dagger \, \psi_R \, \hat{\jmath} \;+\; \hat{\jmath}\,\widetilde{\psi}_R \, \psi_L^* \right) \;+\; m\left( \widetilde{\psi}_L \, \psi_R^* \, \hat{\jmath} \;+\; \hat{\jmath}\,\psi_R^\dagger \, \psi_L \right). \qquad \text{(B.60)} \quad \texttt{\{L-dirac-expan}$$

Caution must be executed with respect to signs when conjugating fermions. It is advisable that the reader should take a moment to understand why the signs in front of the conjugated terms are as they are shown in Eq. (B.60).

Let us begin with the kinetic terms first. Because of Lorentz transformations (and of gauge transformations, if applicable), $\psi_L$ and $\psi_R$ are traditionally written on the right, and their hermitean conjugates — on the left. So it is only straightforward for us to apply operator $\partial_{\psi_L}^\dagger$ from (B.56) on the left. By construction, operators in (B.56) erase the fermions directly neighbouring them. The first kinetic term gives us $i\mathcal{D}\,\psi_L$. What about $\widetilde{\psi}_L \, i\mathcal{D}^*\,\psi_L^*$ ? Equation (B.55) seems to imply that the derivative of $\psi_L^*$ with respect to $\psi_L^\dagger$ maybe non-zero. While this is actually true, the kinetic term for $\psi_L^*$ simply gets filtered out by the projector $P_L$ from (B.56). The same happens to the mass term containing $\psi_L^*$, so that only the first mass term contributes, and we arrive to,

$$i\mathcal{D}\,\psi_L \;-\; m\,\psi_R\,\hat{\jmath} \;=\; 0\,. \qquad \text{(B.61)} \quad \texttt{\{left-dirac\}}$$

Analogously, the derivative with respect to $\psi_R^\dagger$ of right-handed kinetic terms is $i\widetilde{\mathcal{D}}\,\psi_R$.

As for the mass term, one can proceed with it in two ways. First, perhaps the easiest, is to re-write it in the form identical to that of the left-handed mass term. That is, observe that the last bracket in Eq. (B.60) is the "zeroth" quaternionic component of $\hat{\jmath}\,\psi_R^\dagger\,\psi_L$, and as such, is invariant to cyclic permutations of the latter. Using this cyclic symmetry we can pull $\psi_R^\dagger$ to the left,

$$m\left( \widetilde{\psi}_L \, \psi_R^* \, \hat{\jmath} \;+\; \hat{\jmath}\,\psi_R^\dagger \, \psi_L \right) \;=\; m\left( \psi_R^\dagger \, \psi_L \, \hat{\jmath} \;+\; \hat{\jmath}\,\widetilde{\psi}_L \, \psi_R^* \right). \qquad \text{(B.62)}$$

36

The problem of differentiating with respect to $\psi_R^\dagger$ becomes trivial for the first term in the bracket, giving us $m\,\psi_L\,\hat{\jmath}$. The second mass term again gets projected out. Altogether, we have

$$i\,\widetilde{\mathcal{D}}\,\psi_R \;\; + \;\; m\,\psi_L\,\hat{\jmath} \;\; = \;\; 0\,. \tag{B.63} \quad \{\texttt{right-dirac}\}$$

We have mentioned here that the derivative $\partial_{\psi_L^\dagger}\,\psi_L^*$ (along with the corresponding right-handed one) is non-zero. What is it then? Let us see,

$$\partial_{\psi_L^\dagger}\,\psi_L^* \;\; = \;\; \partial_{\psi_L^\dagger}\,(\psi_D\,P_L)^* \;\; = \;\; \partial_{\psi_L^\dagger}\,\psi_D^*\,P_R \;\; = \;\; \frac{1}{2}\,\partial^*\,P_L\,\psi_D^*\,P_R\,. \tag{B.64}$$

Using the first identity in (B.29) (for, $\partial^*$ for $\psi_D^*$ plays the same rôle as $\partial$ does for $q$), we find,

$$\partial_{\psi_L^\dagger}\,\psi_L^* \;\; = \;\; -\,\widetilde{P}_L\,P_R \;\; = \;\; -\,P_R\,. \tag{B.65}$$

It is because of the other factors containing projectors that this derivative did not contribute to the equation of motion (B.61) (and, correspondingly, the analogous right-handed derivative did not contribute to Eq. (B.63)).

The second way of differentiating the mass term with respect to $\psi_R^\dagger$ is to use Eq. (B.56) directly. This would involve applying derivative $\partial^*$ with the help of identities (B.29). The result will of course be the same.

# References

[1] S. L. Adler, New York, USA: Oxford Univ. Pr. (1995) 586 p. (International series of monographs on physics, 88)

[2] J. C. Maxwell, "A Treatise on Electricity and Magnetism," Oxford, Clarendon Press (1873)

[3] D. Finkelstein, J. M. Jauch and D. Speiser, J. Math. Phys. **3**, 207 (1962). doi:10.1063/1.1703794

[4] D. Finkelstein, J. M. Jauch, S. Schminovich and D. Speiser, J. Math. Phys. **4**, 788 (1963). doi:10.1063/1.1724320

[5] J. D. Edmonds, Int. J. Theor. Phys. **6**, 205 (1972). doi:10.1007/BF00672074

[6] J. D. Edmonds, Lett. Nuovo Cim. **7S2**, 398 (1973) [Lett. Nuovo Cim. **7**, 398 (1973)]. doi:10.1007/BF02735143

[7] J. D. Edmonds, Int. J. Theor. Phys. **10**, 115 (1974). doi:10.1007/BF01810397

[8] M. Gunaydin and F. Gursey, J. Math. Phys. **14**, 1651 (1973). doi:10.1063/1.1666240

[9] M. Gunaydin and F. Gursey, Phys. Rev. D **9**, 3387 (1974). doi:10.1103/PhysRevD.9.3387

[10] W. Gough Eur. J. Phys. **7**, 35 (1986). doi:10.1088/0143-0807/7/1/007

[11] W. Gough Eur. J. Phys. **10**, 188 (1989). doi:10.1088/0143-0807/10/3/005

[12] S. De Leo, Int. J. Mod. Phys. A **11**, 3973 (1996) doi:10.1142/S0217751X96001863 [hep-th/9508010].

[13] S. De Leo and P. Rotelli, Mod. Phys. Lett. A **11**, 357 (1996) doi:10.1142/S0217732396000400 [hep-th/9509059].

[14] S. De Leo and W. A. Rodrigues, Jr., Int. J. Theor. Phys. **37**, 1511 (1998) doi:10.1023/A:1026611718277 [hep-th/9806057].

[15] S. De Leo, Found. Phys. Lett. **14**, no. 1, 37 (2001) doi:10.1023/A:1012077227985 [hep-th/0103129].

[16] C. Furey, Phys. Rev. D **86**, 025024 (2012) doi:10.1103/PhysRevD.86.025024 [arXiv:1002.1497 [hep-th]].

[17] C. Furey, JHEP **1410**, 046 (2014) doi:10.1007/JHEP10(2014)046 [arXiv:1405.4601 [hep-th]].

[18] C. Furey, Phys. Lett. B **742**, 195 (2015) doi:10.1016/j.physletb.2015.01.023 [arXiv:1603.04078 [hep-th]].

[19] C. Furey, arXiv:1611.09182 [hep-th].

[20] P. Girard, *Quaternions, Clifford algebras and relativistic physics*, Birkhäuser Verlag (2007).