

A new linear theory of light and matter

John G. Williamson

Quicycle, Scotland, UK

E-mail: quicycle@gmail.com

Abstract. A new theory is proposed with similar form and mechanics to the Dirac equation, but where the rest-mass is introduced in a more fundamental way. The new theory proves to encompass and extend classical electromagnetism, and reduce to it exactly in the field only case. A sharpening of the principle of relativity allows fully relativistic field only solutions, which take on some of the properties more usually associated with quantised states. The extended theory has four new solutions which exhibit a non-zero electric field divergence and a double-covering symmetry. These are necessarily charged, essentially fermionic states, two with positive and two with negative charge, differing by an inner spin symmetry. The new configurations are identified with the spin “up” and spin “down” electron and positron.

1. Introduction

In this paper a new theory will be outlined that extends and encompasses classical electromagnetism, uses some of the form and machinery of relativistic quantum mechanics, but introduces mass in a more sophisticated way than in the Dirac theory. The resultant framework proves to describe not only classical light, but also some aspects more usually associated with its quantum nature. The extension allows new solutions of coupled light and matter. These are necessarily charged and have internal spin. The new solutions are identified with the “elementary” electron and positron. To allow the new developments, a mathematics has been developed which parallels the relativistic structure of physical space and time as closely as possible. The paper has been written in such a way that the physics may be followed without a complete grasp of the mathematics. For those more familiar with the usual Maxwell equations, or on a first reading of the paper, it may be found convenient to skip to section 4 where the extended equation is written in a more familiar 3-dimensional algebra, valid in a given Lorentz frame, from around Eq. (8). One could then skip sections 5 and 6 and resume in section 7 where the proposed mechanism allowing the self-confinement of light and matter into particle-like configurations is discussed.

2. Outline of the axioms of the new theory

The basis taken is of root-energy in space-time forms. Space and time are introduced as unit vectors, which may take any magnitude. The unit basis vector in time is denoted α_0 , and the three right-handed orthonormal unit vectors in space α_1 , α_2 and α_3 respectively. These are unit line-elements. A vector is written with a single index as, for example, Ξ_μ . In a given frame, a 4-vector is written as:

$$\Xi_\mu = \xi_\mu \alpha_\mu = \xi_0 \alpha_0 + \xi_1 \alpha_1 + \xi_2 \alpha_2 + \xi_3 \alpha_3 \quad (1)$$

The α_μ are a set of four fundamental linearly independent basis vector unit elements of time and space. The coefficients ξ_μ are real number magnitudes. Thus far this is a purely conventional 4-vector algebra. For the spatial components, the right-handed ordered triple (1, 2, 3) will be used for the general case. For the case of light, Cartesian (x, y, z) or cylindrical, (r, ϕ, z) may be most appropriate. For particles such as the electron one may wish to use spherical co-ordinates in normal space (r, θ, ϕ) or toroidal co-ordinates in field space (ρ, θ, ϕ) . The base theory is developed in the first instance in the most familiar and recognisable of these, the Cartesian system. Here and in the sequel, Greek indices run from 0 to 3 and Roman from 1 to 3, with 0 representing the time “direction” in individual components in any co-ordinate system so that $\xi_\mu \alpha_\mu = \xi_0 \alpha_0 + \xi_i \alpha_i$. Natural units are used, though all ξ quantities may be most appropriately thought of as being in square root-energy density units, $\sqrt{Jm^{-3}}$. From the four unit lines, α_μ , are generated, by a “product” representing a physical overlap, twelve other space-time geometrical constructs: unit planes, volumes, a hypervolume and, most importantly for the simplest extension of the Maxwell equations to be discussed here, a unit “point”. The concept here is that of a point as opposed to a line or a plane, not of an arbitrarily small thing. In particular, like a directed line or plane element, it may take any magnitude and either sign. Because it is the resultant of a product of a 4-vector with itself, it represents a Lorentz invariant density, rather than the next to nothing often associated with a “mathematical point”. It is at this point that the first departure is made from the conventional development. To make connections with earlier work, and because of its effect in conjunction with the electric field of turning the direction of the flow of electromagnetic momentum the mass-energy associated with this physical “direction” will be called here the “pivot” and denoted with the symbol α_P . The algebra is designed such that the product of any unit line, plane, volume or hypervolume with itself is either α_P or $-\alpha_P$.

The properties of the unit scalar element, the pivot, are consistent with relativity and experiment where the square of 4-vectors are Lorentz invariant scalars. A space time metric (+, -, -, -) is adopted. The unit scalar is taken to be idempotent, such that $\alpha_P^2 = +\alpha_P$, so that a squared scalar root-energy-density is an invariant scalar energy density. The three spatial components α_1, α_2 and α_3 are defined to square to $-\alpha_P$ and the temporal component α_0 to $+\alpha_P$. Explicitly: $\alpha_0^2 = +\alpha_P$, $\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \alpha_i^2 = -\alpha_P$. This leads to the square of 4-vector objects being Lorentz invariant scalar densities, as they should be. Explicitly: $\Xi_\mu^2 = \alpha_P(\xi_0^2 - \xi_1^2 - \xi_2^2 - \xi_3^2)$. Because there are four unit lines there are six unit planes, properly denoted bi-vectors. These may be derived from the basis unit lines by multiplying or dividing mutually perpendicular pairs such that a unit (1,2) plane element is, for example, $\alpha_1 \alpha_2 = \alpha_{12} = -\alpha_{21}$. These unit plane elements then inherit the property of squaring to an invariant scalar e.g. $\alpha_{12}^2 = \alpha_1 \alpha_2 \alpha_1 \alpha_2 = -\alpha_P$. Explicitly these six are: $\alpha_{23}, \alpha_{31}, \alpha_{12}, \alpha_{01}, \alpha_{02}, \alpha_{03}$. The general bivector form is written with two indices as $\Xi_{\mu\nu}$. In a given frame this multi-vector may be decomposed in components as $\Xi_{\mu\nu} = \xi_{0i} \alpha_{0i} + \xi_{jk} \alpha_{jk} = \Xi_{0i} + \Xi_{jk} = \xi_{01} \alpha_{01} + \xi_{02} \alpha_{02} + \xi_{03} \alpha_{03} + \xi_{23} \alpha_{23} + \xi_{31} \alpha_{31} + \xi_{12} \alpha_{12}$. Here, the $\xi_{\mu\nu}$ are real number coefficients and the $\alpha_{\mu\nu}$ are the six linearly independent unit plane elements. Differing indices imply differing values (e.g. in $\Xi_{\mu\nu}$, $\mu \neq \nu$). Same indices imply a summation convention. Again, these quantities inherit the property that their square is a unit scalar. To jump forwards slightly and to give a physical example to be derived later in the section on the correspondence with the Maxwell equations, these components of the bi-vectors transform under a Lorentz transformation as do the (1,2,3,) components of the electric $(-\xi_{0i})$ and magnetic $(-\xi_{jk})$ field respectively[1]. In the case of division, one must take care of the distinction between quantities divided by or divided into other quantities. Here, unless otherwise stated, the convention is adopted that $\frac{\alpha_a}{\alpha_b} = (\frac{1}{\alpha_b})\alpha_a$. The rationale for this choice is that this gives signs corresponding to the conventional ones for the 4-vector differentials at the level of the Maxwell equations, and where the vector operator is taken to act to the “right”. It is worth noting that many conventions of operator ordering, the sign of base quantities such as the electron charge and the handedness of the spatial basis, are human inventions. Given

the number of permutations for these choices, it is unlikely that humans have chosen the same conventions as has nature in every respect, although one may be lucky. Note that the magnitudes which have been introduced parallel those of the probability density in quantum mechanics, or the field in electromagnetism in that they must be squared and integrated over to represent a probability (or energy) density. They may be viewed as representing square-root density magnitudes: this is as it should be, it is the square of the field that is proportional to an energy density, not the field itself. Here, all the elements of Ξ are taken as fundamentally square-root energy densities. The 4 tri-vectors (the dual of the vectors) represent unit volume elements in space and time. Explicitly these are the outward directed spatial volume element, the hedgehog, denoted α_{123} (spines outward, the inward directed element $-\alpha_{123}$ is a less comfortable image) and the directed space-time “volumes”, α_{023} , α_{031} , α_{012} . The latter three transform as a momentum density multiplied by a perpendicular unit vector: that is they transform as the components of an intrinsic angular momentum density. These tri-vectors have the general multi-vector form $\Xi_{\mu\nu\rho} = \xi_{023}\alpha_{023} + \xi_{031}\alpha_{031} + \xi_{012}\alpha_{012} + \xi_{123}\alpha_{123}$. Finally, there is a single quadri-vector, representing an outward directed 4-volume element, the quedgehog, α_{0123} which, just as the scalar pivot, is invariant under a Lorentz transformation but, analogously with the unit imaginary i , squares to the negative unit scalar element $\alpha_{0123}^2 = -\alpha_P$. A quantity with this form is written $\Xi_{0123} = \xi_{0123}\alpha_{0123}$. In total, there are then sixteen linearly independent unit space-time forms. Four “lines”, six derived planes, four volumes a hypervolume and a derived “point direction” α_P . In this metric, six elements square to the positive scalar unity α_P : α_0 , α_{01} , α_{02} , α_{03} , α_{123} and α_P itself. The remaining ten square to the negative scalar unity $-\alpha_P$: α_1 , α_2 , α_3 , α_{23} , α_{31} , α_{12} , α_{023} , α_{031} , α_{012} and α_{0123} . The algebra so defined has useful sub-algebras. The even subset is closed under multiplications. The set $[\alpha_P, \alpha_{23}, \alpha_{31}, \alpha_{12}]$ is isomorphic to the quaternions. The set $[\alpha_P, \alpha_{0123}]$ is isomorphic to complex numbers. If one were to make the identification $\alpha_P^2 = \alpha_P = 1$, the resultant would be the Clifford algebra $Cl(1,3)$ [2, 3] or the space-time algebra championed by Hestenes as the STA in internal models of the electron structure [4]. In contrast to some other work, the functions of magnitude, sign and form are separated explicitly in the formalism, according to the precepts of “absolute relativity” discussed in earlier work[5, 6, 7]. Briefly, absolute relativity requires that, in the description of 4-dimensional solutions or equations, every derived quantity must be properly related to the physical properties of space and time. Absolute relativity is implemented here by demanding strictly that no quantity in any 4-dimensional solution or equation may appear without the proper, relativistic unit element representing its space-time form. The new formalism applies this rigorously not only, as is usual, to vectors, bi-vectors and so on, but also to differentials, power series and even to the arguments of functions. It is this latter property that allows properly fully relativistic wave-functions to be written down[6]. It should be noted that, for example, this does not allow the use of the unit imaginary i squaring to the real number -1, unless this can be properly associated with a physical dimension in space, time or products of space and time. The closest analog to the usual unit imaginary i is the quedgehog since the sub-algebra $[\alpha_P, \alpha_{0123}]$ is isomorphic to that of $[1, i]$ in the simpler complex system. In the full algebra, however, α_{0123} is in no way equivalent to the unit imaginary.

A general multi-vector $\Xi_{\mathcal{G}}$, may be defined over each of the sixteen multi-vector quantities defined above with a scalar term Ξ_P , a vector term Ξ_μ , space-time and space-space bi-vector forms Ξ_{0i} and Ξ_{jk} such that $\Xi_{\mu\nu} = \Xi_{0i} + \Xi_{jk}$, a tri-vector term $\Xi_{\mu\nu\rho}$ and an eventual quadri-vector term Ξ_{0123} . Explicitly one may define a sixteen component multi-vector:

$$\begin{aligned} \Xi_{\mathcal{G}} = \xi_{\mathcal{G}}\alpha_{\mathcal{G}} = & \xi_P\alpha_P + \xi_0\alpha_0 + \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3 + \xi_{01}\alpha_{01} + \xi_{02}\alpha_{02} + \xi_{03}\alpha_{03} + \\ & \xi_{23}\alpha_{23} + \xi_{31}\alpha_{31} + \xi_{12}\alpha_{12} + \xi_{023}\alpha_{023} + \xi_{031}\alpha_{031} + \xi_{012}\alpha_{012} + \xi_{123}\alpha_{123} + \xi_{0123}\alpha_{0123} \end{aligned} \quad (2)$$

Where the subscript \mathcal{G} takes any or all of the combinations associated with the sixteen linearly independent unit elements derived from multiplication amongst the basis elements α_i ,

enumerated above, and the $\xi_{\mathcal{G}}$ are all real number coefficients (some of which may be zero) representing terms transforming as square-root energy densities.

3. General linear equation for Ξ

Using the formalism outlined above, a 4-vector 4-differential is defined as:

$$\mathcal{D}_\mu = \frac{\partial}{\alpha_\mu \partial x_\mu} = \frac{\partial_\mu}{\alpha_\mu} = \alpha_0 \partial_0 - \alpha_i \partial_i = \alpha_0 \partial_0 - \alpha_1 \partial_1 - \alpha_2 \partial_2 - \alpha_3 \partial_3 \quad (3)$$

Absolute relativity is imposed in that the differential elements contain an implicit quotient of the unit vectors. This results in a change of sign of the spatial components due to the fact that the three spatial vectors α_1, α_2 and α_3 square to the negative scalar unity: $\alpha_i^2 = -\alpha_P$ and $\alpha_i \alpha_P = \alpha_i$. The differential operator acting on any multi-vector results in a change of form by one index, from the “odd” set to the “even” set or vice-versa. In the case of the same indices cancelling this represents a transformation inwards, for example from a plane to a line. If a new index is introduced, this represents a transformation outwards, for example from a plane to a volume. Taking the 4-vector differential, defined by Eq. (3), of Eq. (2) and grouping the scalars in terms of the resultant unit α ’s one obtains by operating from the left and setting this equal to some function (zero in the simplest case) a set of 16 coupled differential equations:

$$\begin{aligned} \mathcal{D}_\mu \Xi_{\mathcal{G}} = \mathcal{F}_{\mathcal{G}} = & \\ & \alpha_0(\partial_0 \xi_P - \partial_1 \xi_{01} - \partial_2 \xi_{02} - \partial_3 \xi_{03}) + \\ & \alpha_{123}(\partial_0 \xi_{0123} - \partial_1 \xi_{23} - \partial_2 \xi_{31} - \partial_3 \xi_{12}) + \\ & \alpha_1(-\partial_1 \xi_P + \partial_0 \xi_{01} - \partial_2 \xi_{12} + \partial_3 \xi_{31}) + \\ & \alpha_2(-\partial_2 \xi_P + \partial_0 \xi_{02} + \partial_1 \xi_{12} - \partial_3 \xi_{23}) + \\ & \alpha_3(-\partial_3 \xi_P + \partial_0 \xi_{03} - \partial_1 \xi_{31} + \partial_2 \xi_{23}) + \\ & \alpha_{023}(+\partial_0 \xi_{23} - \partial_1 \xi_{0123} + \partial_2 \xi_{03} - \partial_3 \xi_{02}) + \\ & \alpha_{031}(+\partial_0 \xi_{31} - \partial_2 \xi_{0123} - \partial_1 \xi_{03} + \partial_3 \xi_{01}) + \\ & \alpha_{012}(+\partial_0 \xi_{12} - \partial_3 \xi_{0123} + \partial_1 \xi_{02} - \partial_2 \xi_{01}) + \\ & \alpha_P(\partial_0 \xi_0 + \partial_1 \xi_1 + \partial_2 \xi_2 + \partial_3 \xi_3) + \\ & \alpha_{0123}(\partial_0 \xi_{123} + \partial_1 \xi_{023} + \partial_2 \xi_{031} + \partial_3 \xi_{012}) + \\ & \alpha_{01}(+\partial_0 \xi_1 + \partial_1 \xi_0 + \partial_2 \xi_{012} - \partial_3 \xi_{031}) + \\ & \alpha_{02}(+\partial_0 \xi_2 + \partial_2 \xi_0 - \partial_1 \xi_{012} + \partial_3 \xi_{023}) + \\ & \alpha_{03}(+\partial_0 \xi_3 + \partial_3 \xi_0 + \partial_1 \xi_{031} - \partial_2 \xi_{023}) + \\ & \alpha_{23}(+\partial_0 \xi_{023} + \partial_1 \xi_{123} - \partial_2 \xi_3 + \partial_3 \xi_2) + \\ & \alpha_{31}(+\partial_0 \xi_{031} + \partial_2 \xi_{123} + \partial_1 \xi_3 - \partial_3 \xi_1) + \\ & \alpha_{12}(+\partial_0 \xi_{012} + \partial_3 \xi_{123} - \partial_1 \xi_2 + \partial_2 \xi_1) \end{aligned} \quad (4)$$

Where in the case corresponding most closely to the usual formulation of the Maxwell equations, $\mathcal{F} = -\Xi_{\mu\nu}$ as discussed below.

4. Correspondence with the Maxwell equations

It may not be immediately apparent, but the above equations encompass the structure of not only the free-space Maxwell equations, but also the constitutive equations of the field including the Lorentz gauge condition. To aid physical thinking, the single-index elements α_i represent a kind of current, but a root-mass current rather than an electrical current. These are best thought of in conventional terms as being related to the “vector potential”. Everything else may be seen

as either a derivative or product of 4-vector terms (in both the mathematical and the more conventional senses of the words). The two index terms $\alpha_{\mu\nu}$ may then be identified with the six elements of the electric and magnetic field, the three index terms a kind of electromagnetic spin density and the pivot and 4-index quadri-vector terms, α_P and α_{0123} , (root) rest-mass components. Note that the patterns arising in the full 4-differential of a general multi-vector, $\mathcal{D}_\mu \Xi_{\mathcal{G}}$ in Eq. (4) are similar in form to those of the time derivative, div, grad and curl familiar from three-dimensional algebras. Look, for example, at the last term in α_{12} of Eq. (4). The first term is a time derivative, the second looks like a component of a divergence (as is matched by the two terms above it). The final pair looks the same as the negative of the curl component of the three-vector (ξ_1, ξ_2, ξ_3) , the negative arising from the negative sign in the definition of the 4-vector derivative in Eq. (3). In making the connection with classical electromagnetism, this will prove to correspond to just the (negative of the) “3” component of the magnetic field. This may be seen as it is part of the derivative of the 4-vector (usually identified with a vector potential) as is illustrated below. If one looks further than the set in α_{ij} to the other three-component sets of α_i , α_{0i} and α_{0ij} above, one sees the same pattern of time derivative, divergence and curl. It is this remarkable correspondence of the 4-dimensional algebra with the more venerable 3-dimensional forms which may be the reason the universe around us appears so three dimensional: in observing an everyday item such as the object on which this article rests, for each space-time point one has a triple of current, α_i , a triple of electric field, α_{0i} , a triple of magnetic field, α_{ij} , and a triple of angular momentum, α_{0ij} , all superimposed and all behaving under rotations and translations in a 3-D fashion[1]. That is, in this formalism the equation:

$$\mathcal{D}_\mu \Xi_{\mathcal{G}} = \mathcal{D}_\mu (\Xi_\mu + \Xi_{\mu\nu}) = -\Xi_{\mu\nu}, \quad (5)$$

with all other terms implicitly zero, encompasses not only all four Maxwell-Heaviside equations, but also the Lorenz gauge condition and the constitutive equations of the fields from the vector potential, as is now shown.

To understand how this appears in the shadow-world of 3-D in a given frame, and hence write equations (4) and (5) in the more conventional form, one needs a 3-vector notation for the 3-component objects which does not contain the 4-dimensional unit elements, though they may contain implicitly the 3-D unit elements in a given frame. Define a 3-vector derivative as $\vec{\nabla} = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$, where the over-arrow denotes such a three-vector set. For example, the 3 components of Ξ_{ij} may then be written as $\vec{\xi}_{ij} = (\xi_{23}, \xi_{31}, \xi_{12})$. $\vec{\nabla}$ and $\vec{\xi}_{ij}$ here are three-component row or column vectors without their proper 4-D unit vector elements, as distinct from Ξ_{ij} , which does contain them. Defining a further set of the three spatial bi-vectors $\vec{\alpha}_{ij} = (\alpha_{23}, \alpha_{31}, \alpha_{12})$ one may define $\vec{\xi}_{ij} \cdot \vec{\alpha}_{ij} = \Xi_{ij}$. First, write the 4-vector potential and the 4-vector differential as:

$$\Xi_\mu = \xi_0 \alpha_0 + \vec{\xi}_i \cdot \vec{\alpha}_i, \quad \mathcal{D}_\mu = \alpha_0 \partial_0 - \alpha_i \cdot \vec{\nabla} \quad (6)$$

The subset of the 16 ($= 1 + 3 + 3 \cdot 2 + 3 \cdot 2$) terms of the 4-derivative \mathcal{D}_μ of the 4-vector Ξ_μ is:

$$\mathcal{D}_\mu \Xi_\mu = \alpha_P (\partial_0 \xi_0 + \vec{\nabla} \cdot \vec{\xi}_i) + \alpha_{0i} (\partial_0 \xi_i + \vec{\nabla} \xi_0) - \alpha_{ij} (\vec{\nabla} \times \vec{\xi}_i) \quad (7)$$

which is the sum of a scalar part Ξ_P , which expresses a gauge condition and a bi-vector part $\Xi_{\mu\nu}$ which, if the vector is associated with a 4-vector potential, has the form and conventional content of the electromagnetic fields. Setting the rest mass-energy term in α_P , which is then the 4-divergence of the 4-vector potential, to zero as in Eq. (4) manifestly corresponds to choosing the Lorenz gauge in this formalism. Other gauges may be chosen by setting this to other scalar constants. Conventionally[8], if one identifies the vector part with a 4-vector potential such that $\vec{A} = \vec{\xi}_i$ then $\vec{B} = +\vec{\nabla} \times \vec{A} = -\vec{\xi}_{ij}$ in Eq. (7), and $\vec{E} = -\partial_0 A_i - \vec{\nabla} A_0 = -\vec{\xi}_{0i}$, such that

$F_{\mu\nu} = \vec{E} + \vec{B} = -\vec{\xi}_{0i} - \vec{\xi}_{ij}$. Defining further terms, a 4-vector such that $A_0 + \vec{A} = \xi_0 + \vec{\xi}_i$, a 4-tri-vector $T_0 + \vec{T} = \xi_{123} + \xi_{0ij}$, a pivot component magnitude $P = \xi_P$ and the quadri-vector quedgehog magnitude $Q = \xi_{0123}$, setting the 4-differential to zero in Eq. (4) such that $\mathcal{D}_\mu \Xi_{\mathcal{G}} = 0_{\mathcal{G}}$ one obtains:

$$\begin{aligned}
& \alpha_0(\vec{\nabla} \cdot \vec{E} + \partial_0 P) + \\
& \alpha_{123}(\vec{\nabla} \cdot \vec{B} + \partial_0 Q) + \\
& \alpha_i(-\partial_0 \vec{E} - \vec{\nabla} P + \vec{\nabla} \times \vec{B}) + \\
& \alpha_{0jk}(-\partial_0 \vec{B} - \vec{\nabla} Q - \vec{\nabla} \times \vec{E}) + \\
& \alpha_p(+\vec{\nabla} \cdot \vec{A} + \partial_0 A_0) + \\
& \alpha_{0i}(+\partial_0 \vec{A} + \vec{\nabla} A_0 + \vec{\nabla} \times \vec{T}) + \\
& \alpha_{jk}(+\partial_0 \vec{T} + \vec{\nabla} T_0 - \vec{\nabla} \times \vec{A}) + \\
& \alpha_{0123}(+\vec{\nabla} \cdot \vec{T} + \partial_0 T_0) = 0_{\mathcal{G}}
\end{aligned} \tag{8}$$

This is not a new result in itself, but a particular representation of a subset of the more general eq. (4) in a particular frame, expressed in the conventional quantities \vec{E} and \vec{B} and in a 3-dimensional algebra. Taking the first four equations, and setting the mass-energy terms to zero so that only the fields remain, dropping the proper 4-dimensional unit elements, by inspection, one has exactly the familiar free-space Maxwell equations. Note that setting the 4-derivative of the fields to zero gives all four free-space Maxwell equations, not just two as in the standard tensor approach[8]. The Maxwell equations, derived largely from careful experiment, have been known for over a century and it has long been known that approaches using Clifford algebras lead to a more compact form of the Maxwell equations[3, 9]. It is not the aim here to merely explore another derivation of classical electromagnetism, but to further the understand the underlying nature of light and material particles. Though the bi-vector subset precisely parallels the free-space Maxwell equations in a particular frame, much of the developments follow from retaining the fully relativistic, frame-independent algebra of eq. (4). In particular, it is the aim in the following to derive charge, understand its underlying nature and hence underpin quantum electrodynamics, rather than simply introduce charge by hand.

5. Relation to the Dirac equation

In natural units, the Dirac equation of relativistic quantum mechanics is often written:

$$(i\rlap{\not{D}} - m)\psi = 0 \tag{9}$$

Here, $\rlap{\not{D}}$ is identical to the operator \mathcal{D}_μ , in that both contain the 4-vector differential and the Dirac γ operators, isomorphic within the sub-algebra to the α operators in \mathcal{D}_μ . Eq. (4) is similar to the Dirac equation in that both are first-order 4-vector-differential equations and both introduce mass terms. They differ, however, in several respects. Eq. (9) introduces the mass as a simple, inert “lump” whereas Eq. (4) brings it in as a dynamical root-mass density in $\Xi_{\mathcal{G}}$ on the same physical basis as the fields. Eq. (9) introduces the unit imaginary where the basis of Eq. (4) remains space, time, field and so on. Eq. (9) operates on a function ψ to be determined whereas Eq. (4) operates on the physical fields and the dynamical root-mass. In summary, the generalised Maxwell equation including root mass density, $\mathcal{D}_\mu \Xi_{\mathcal{G}} = 0$ (Eq. (4)), may be seen as a generalisation of the Dirac equation, where the mass is treated on a more sophisticated way, as a dynamical quantity on the same footing as the fields, and the wave-function is more closely related to the solutions of the Maxwell equations in that they refer more closely to the observed physical quantities such as field, current and angular momentum. Given the similarity of the equations one might, nonetheless, expect similar solutions and this is the case, as is now explored.

6. Solutions to the Maxwell equations constrained by “absolute relativity”

Now, as is well known, not all solutions to the Maxwell equations are physical. In particular, most real photons follow the simple quantisation rule that their energy \mathcal{E} is characterised by the frequency ω alone, such that $\mathcal{E} = \hbar\omega$. For a photon both the local scales of length and time (its wavelength) and the change in field strength must be taken into account in determining its integrated energy. The photon frequency and hence energy change in the relativistic Doppler effect is given by the Doppler factor.

$$R = \sqrt{\frac{1+\beta}{1-\beta}} = \gamma(1+\beta), \frac{1}{R} = \sqrt{\frac{1-\beta}{1+\beta}} = \gamma(1-\beta) \quad (10)$$

With:

$$\beta = \frac{R^2 - 1}{R^2 + 1}, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{1}{2}\left(R + \frac{1}{R}\right) \quad (11)$$

Provided the magnitudes of the electric and magnetic field components are equal (as they are for propagating free-space electromagnetic waves in general and for photons in particular), so that $|E| = |B|$, in root-energy terms they transform relativistically as:

$$E' = \gamma(E + \beta B) = RE (= RB), B' = \gamma(B + \beta E) = RB (= RE) \quad (12)$$

That is, for light, the resultant fields transform in the same way as does the frequency and energy: linearly with R . This is where the power of the fully-relativistic algebra is essential: to write a wave-function consistent with absolute relativity, one must include the proper form of space and time in the exponent. Such a function represents a wave only if an additional factor transforming both space and time to factors which square to $-\alpha_P$ [6, 7]. For a wave function in the “3” (z) direction, such a factor is the unit angular momentum about that direction α_{012} [6]. Such a wave-function is

$$\Psi = \mathcal{X}_0 e^{\alpha_{012}(\alpha_3 k z - \alpha_0 \omega t)} = \mathcal{X}_0 e^{\alpha_{0123} k z} e^{-\alpha_{12} \omega t} \quad (13)$$

Note that, in the exponent, both the “z” component of space and the time differ by their proper vector forms α_3 and α_0 respectively, as they must under the constraint of absolute relativity. This is in contrast to conventional wave-functions, where the phase appears as a scalar factor [8, 3]. This function, while it represents a wave and is a solution to Eq. (4), is not a pure-field wave, as the wave has four components, two transforming as a root-mass (α_P and α_{0123}) and two transforming as a field (α_{12} and α_{03}). It is a mass-field wave-function. It is not a photon wave function, but is a candidate wave-function for particles with both field and rest-mass, such as the electron. Though this has more internal complexity than a complex wave-function it may lead to the same results, for example for a pure-field configuration, as will be discussed around Eq. (15) below. More importantly, it has a remarkable property that makes it a superb candidate for coupling to light and emitting or absorbing electromagnetic radiation: forming a product of this function with a strongly restricted set of initial pure-field configurations yields a resultant propagating pure field which is a solution to the Maxwell equation alone. That set of restricted field configurations corresponds exactly to those observed in physical photons. Taking a product with an initial field (an overlap) where \vec{E} is perpendicular to \vec{B} and of arbitrary but equal magnitude leads to an additional pure-field rest-massless wave-function for the light-speed case where $k = \omega$, propagating in the proper direction perpendicular to both \vec{E} and \vec{B} . One is led to a wave-function incorporating this and the R factor above into both pre-factor (transforms the field correctly) and exponent (transforms the frequency correctly), such that the

whole function transforms properly relativistically between any inertial frame, governed by the single energy and frequency scale parameter R . Such a function remains massless in all frames and the integral energy, frequency and field scale with R . What emerges is a fully relativistic wave-function representing the same photon in each and every inertial frame. Such a function may be written in Cartesian (or cylindrical) components as:

$$\Xi_\gamma = \mathcal{H}_0 R (\alpha_{01} + \alpha_{31}) e^{R\omega\alpha_{012}(\alpha_3 z - \alpha_0 t)} \quad (14)$$

Where \mathcal{H}_0 is a constant relating the function to energy density units, related to \hbar , or through the relation between spin and charge calculated in earlier work[10] to the elementary charge q . This is proposed as a candidate for the detailed physical mechanism by which a mass-field wave-function like eq. (13) transforms to a quantum of electromagnetic energy subsequently transmitted through space and time as in eq. (14). Note that the product R in the exponent affects only the rate of change of phase. In the pre-factor it represents the change in field strength and the total photon integral energy under a Lorentz transformation[6]. If this function is a solution in any frame, then it is also a solution in any other frame: the field, energy and exponent scale linearly with R and hence the local ruler and clock scale reciprocally. Taking $R = 1$ for a specific photon of frequency ω_1 , for a blue shift, the photon shrinks as $\frac{1}{R}$, the field increases as R and hence the energy density in the field increases as R^2 , leading to an overall increase in integrated energy which scales as R . The integral energy is agnostic to the lateral distribution as field decreases linearly hence energy density quadratical, but the lateral area also increases quadratically, cancelling this contribution. The wavetrain retains the same number of phase cycles, but the relativistic change of length ensures $\mathcal{E} = R\hbar\omega_1 = \hbar\omega$ remains satisfied. That is, for a given \mathcal{H}_0 , all wave functions of the form of eq. (14) have the same, quantised, angular momentum. As R tends to zero the light is redshifted to oblivion, the frequency goes to zero the energy to zero and the field to zero, as is consistent for a zero rest-mass particle. Conversely as R tends to large values, frequency, field and integral energy all increase linearly with R . This parallels precisely that which is observed in nature for the physical photon. Though the energy density varies as a square of the field, the linearity between energy and the frequency and the local field is maintained because the relativistic “box” over which one must integrate the total energy also transforms, scaling precisely with the photon wavelength[1]. The number of oscillations, the total change of phase, is a constant for a given photon regardless of which frame it is viewed from. The resulting change of the scale of space and time, the reference “rulers” and “clocks” is precisely that required to maintain the linearity of both energy and field. This may then be viewed as being the physical reason for special relativity[6]: the transformations of space and time are required to effect the observed linearity of both energy and field in light as expressed by eq. (10). Though Eq. (14) looks unconventional, it may be readily expanded in any particular frame yielding textbook results. On pre-multiplying eq. (13) by a field corresponding to that of a physical photon the mass terms cancel. One obtains:

$$\Xi_\gamma = \mathcal{H}_0 R [(\alpha_{01} + \alpha_{31}) \cos R(kz - \omega t) + (\alpha_{23} - \alpha_{02}) \sin R(kz - \omega t)] \quad (15)$$

This describes a textbook example of an electric (α_{0i}) and a magnetic (α_{ij}) field rotating in time in a plane perpendicular to the direction of momentum transport and transforming in space from magnetic field to electric field and vice-versa, just as is required by the Maxwell equations. Further the new formalism is not just descriptive, it is proscriptive. For example, swapping the sign of either field requires a reversal of the propagation direction for a solution, as is the case physically. Also note that, just as is the case for the physical photon, any transverse terms will not propagate, as these square to $+\alpha_P$. That the application of the principle of absolute relativity leads to a relativistic linearly propagating, laterally confined, photon-like wave-function of this form is the second result of this paper.

7. An extension of the Maxwell equations to describe charged, spin half particles

Consider the subset of Eq. (4), including the bi-vector field and the scalar pivot alone: the equation $\mathcal{D}_\mu(\Xi_{\mu\nu} + \Xi_P) = 0$. This, the simplest possible extension of the free-space Maxwell equations to include the root-mass-energy term, the pivot, brings in a crucial new feature into the theory: a term allowing electro-pivot-magnetism to confine itself. The relevant extension are illustrated in the first four terms of Eq. (8) these reduce to the free space Maxwell equations for the case where both P and Q are zero. In the case where Q remains zero, but one brings in a non-zero root-mass P , one has extra terms corresponding to charge and current. Indeed, if the time rate of change of root-mass is non-zero so that there is a mass-energy exchange, then the charge, as expressed by the $\vec{\nabla} \cdot \vec{E}$ term, is necessarily non-zero, such that $\partial_0 P = -\vec{\nabla} \cdot \vec{E}$. It is interesting to note that a negative divergence implies the sign of the pivot should be positive. The pivot, being a root energy, can take either sign, with its square, corresponding to an energy density, being positive definite in either case. If pivot exchange is the proper origin of charge, however, it is quite satisfying that the conventional choice of the negative sign of the electron leads to the underlying elementary root-mass being positive here. Note that pivot is not charge: pivot (mass-energy) exchange is at the root of charge. In the simplest extension here it is the time rate of change of pivot that is proportional to the charge.

The way the pivot affects the dynamics of the electromagnetic field is now investigated. Conventionally, the energy density and the momentum density (the Poynting vector) is obtained from the product of field and its conjugate in a similar way to that in which the probability density in quantum mechanics is given by $\Psi\Psi^*$. Consider a system of field and root-mass as $\Xi_{\mathcal{G}} = \Xi_{\mu\nu} + \Xi_P$ and consider the Hermitian conjugate set as $\Xi_{\mathcal{G}}^\dagger = \Xi_{\mu\nu}^\dagger + \Xi_P^\dagger$. Multiplying these out then gives the generalised energy-momentum density $\mathcal{P} = \frac{1}{2}\Xi_{\mathcal{G}}\Xi_{\mathcal{G}}^\dagger$ for this limited system. One obtains:

$$\mathcal{P} = \frac{1}{2}\alpha_P(\vec{E}^2 + \vec{B}^2 + P^2) + \alpha_{0i}(\vec{E} \times \vec{B} + P\vec{E}) \quad (16)$$

Clearly for the case $P = 0$ one has the usual expression for the electromagnetic energy density $\frac{1}{2}(\vec{E}^2 + \vec{B}^2)$ and for the momentum density (the Poynting vector) $\vec{E} \times \vec{B}$ as expected, but with their proper forms of rest-mass and momentum density. The new feature for $P \neq 0$ is the emergence of an extra term in the rest mass-energy density ($\frac{1}{2}P^2$) and an extra term in the momentum density perpendicular to the Poynting vector ($P\vec{E}$). It is this term that leads to the self-confinement of the new particle-like solutions. The rest-mass $\frac{1}{2}P^2$ term contributes to the total mass-energy of the system in the same way as do the elements of the electromagnetic field, or as does the wave function squared in ordinary quantum mechanics. The effect of the $P\vec{E}$ term is that a component of the momentum appears in the direction of the electric field, and hence perpendicular to the Poynting vector, as is clear from Eq. (16). The resultant is that, within a rest-massive particle, the momentum flow turns in the plane of the electric field and Poynting vector. Electromagnetic momentum turns in the presence of pivot. As discussed in earlier work[10, 7], an internal harmony (of a factor of two) between the rate of the twist around the propagation direction described in Eq. (14) and the rate of turn engendered by the electric field-pivot interaction described by Eq. (16) leads to a pair of nested rotations resulting in the spiralling electric field of the photon field being “rectified” such that it becomes everywhere radial outwards or inwards. The resultant is a self-confined charged object. The simplest resonant, coherent system is then a single wavelength wave trapped in a tumbling toroidal double loop configuration in field-momentum space[10]. Because this allows for a maximal field cancellation this will be the lowest energy configuration. Such a configuration, once formed, will be stable, as to perturb it will require energies far larger than the integrated mass-energy of the object. A diagram of the initial formation or annihilation of an electron-positron pair, in this formulation, from two photons, illustrating the field patterns for successive phases is shown in

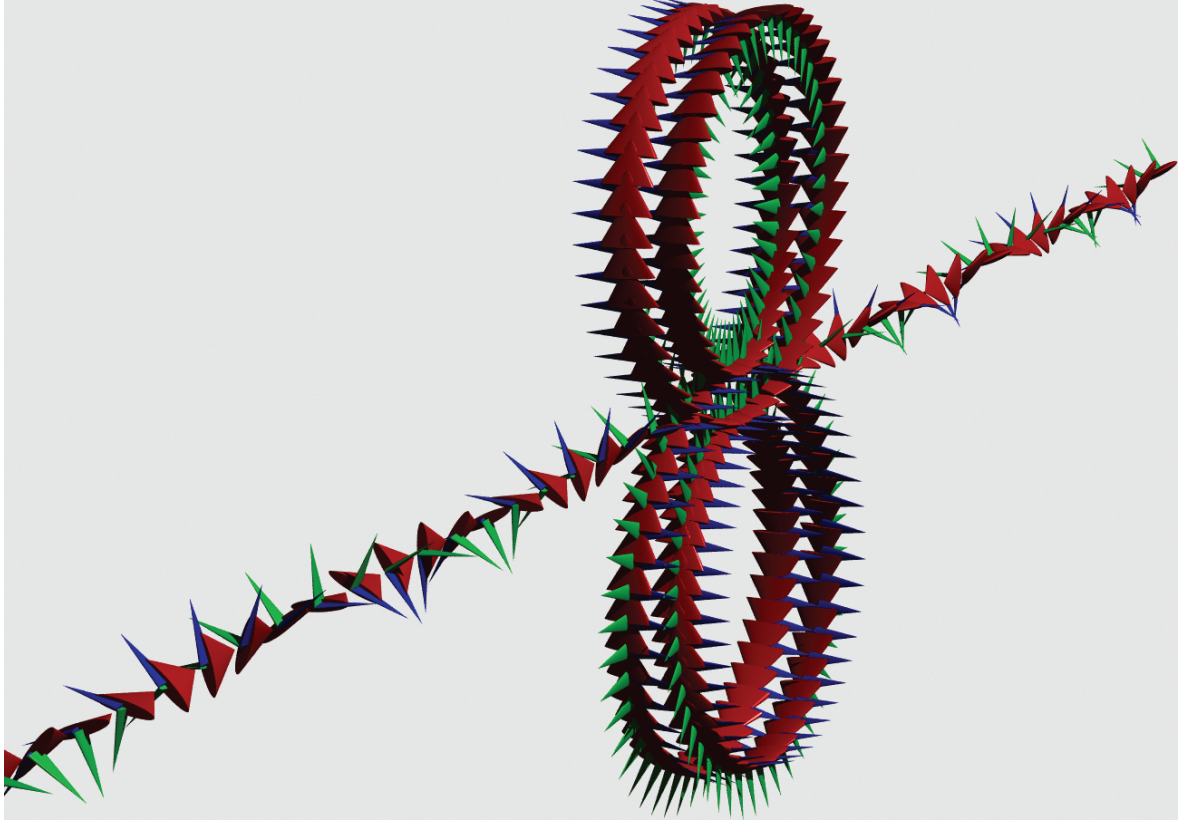


Figure 1. Illustration of the first pair of nested “rotations” in electron-positron vortex pair creation (illustrated) or annihilation from or to a pair of photons at the point of separation. The field distribution for the photons is that of Eq. (14) with the propagation axis the “z” direction in the equation. The initial plane of resultant vortex loops may then be taken to be the y-z plane. The electric field is denoted in green, the magnetic in blue and the momentum density as red arrowheads. For clarity, the incoming photons are shown on a different scale to the electron and positron vortices, which should be 7.5 times smaller than illustrated. The vortices are double-covering single wavelength loops with “radius” a factor of 4π smaller than the photon wavelength. The electron (above) incorporates a pool of positive pivot, the positron (below) a pool of negative pivot. In the process charge, spin and angular momentum are conserved.

Fig. 1. The process of electron-positron decay into two photons (where the rest-mass terms in electron and positron annihilate), or the reverse process of the photo-production of an electron positron pair (where two equal and opposite pools of scalar root-mass are created as well as a circulating internal field configuration) may then be understood as a continuous process such that the set of coupled linear differential eqs. (4) are everywhere satisfied. The electromagnetic field patterns in the initial process of the photo-production of an electron-positron vortex pair (or electron-positron annihilation), at the point where the electron and positron just separate (or where annihilation just begins) are shown in Fig. 1 for the initial photon states (spiralling electric field) and the final electron and positron states (radial electric field). For a complete separation to a free particle-antiparticle pair the initial photon energies must be sufficient to impart enough momentum to both overcome the Coulomb attraction and to cancel to provide sufficient local pivot of both signs to give confinement to the incipient electron and positron vortices. In a stable solution for each separate lepton, each of the elements in the different

space-time forms must then transform between one another smoothly and locally such that each one is continuously replenished as it is destroyed, as described by eq. (4) and as takes place in the wave-functions described in section 6. In the creation or annihilation process, all quantities, angular momentum, charge, current, local root-energy (and hence total integrated energy) must be conserved locally. The net result is that the zero-charge two boson initial state in two cylindrical topologies (the two photons) wraps up into a zero charge final state with two self-sustaining electro-pivot-magnetic vortices with equal and opposite pivot in toroidal topologies (the electron-positron pair) corresponding to the two fermions of a particle-antiparticle pair. Both the initial and final states have equal and opposite spins, charges and scalar components. Conversely, in pair annihilation, the initial fermionic toroidal vortices unwrap to yield a bosonic field-only pair of photons with zero rest mass (pivot), each described by the twisted solution of eq. (14). The proposed relationship between the internal phases for twist and turn extends slightly de Broglie's idea of the "harmony of phases" [11, 12], to a harmony between a set of nested "rotations": the innermost the internal photon phase (a twist-really a field transformation, as in Eq. (14)), the middle the rate of circulation of electromagnetic momentum (a turn- engendered by the pivot-field interaction, as in Eq. (16)) and the outer the rate of rotation of the whole object to give a maximal external field cancellation (a tumble- a real rotation). In a resonant, coherent solution these three phases must be locked with respect to one another to form the lowest energy configuration. The outermost rotation, the tumble, serves to cancel the external magnetic field component to minimise energy. The implicit order of these three nested turns, from the differential, from the gauge and from the external constraints matters, in the sense that they act as "wheels within wheels" for the model electron. For positive pivot, the net result of the collective motion is an electron-like inward directed electric field and for negative pivot an outward directed positron-like field. A single vortex at the point of formation, illustrating the essentially toroidal inner topology in field-space, is shown in Fig. 2, this corresponds to the bottom vortex in Fig. 1. The loops represent the initial (or final) state of the creation (or annihilation) of an electron-positron pair. They have some similarity to photon-loop models proposed by the author and others [10, 13, 14, 15, 16, 17]. The loops have some of the properties of physical spinors, as discussed in earlier work [18, 7]. They do not, however, represent a stable sustainable field distribution for an isolated elementary particle in free space. They are the limit in the case of a very strong (Magnetar strength) external magnetic field. Experimentally, one should easily resolve a loop of the order of $\frac{\lambda_C}{4\pi}$, and this is not observed in experiment for unpolarised electrons. It is also apparent that a simple planar loop will exhibit a huge dipole magnetic field, which is again not observed for a free electron. The electron has a dipole moment only if induced by an external field. The configurations illustrated for the vortices in Fig. 1 are therefore by no means a minimum energy configuration. The outer diagram of Fig. 2, and the loops in Fig. 1 are drawn such that successive phases of the photon wave-function of Eq. 14 are at different points in space, as is the case for a free photon "book-ended" by two charged systems. In reality, as a free electron has nothing to rotate about but itself, the wave-function of the electromagnetic part should be drawn such that successive phases appear at the same point in the space of an outside observer. By symmetry, the resultant is then a perfectly spherical electron, as illustrated in the core of the figure. As the vortex forms, it forms about a single droplet of pivot. Slices through the spherical droplet of pivot, represented by the smoke discs in Fig. 2 are circular. Each disc represents a circular slice through the same spherical drop of pivot at the heart of the particle. In a stable solution, all phases, and hence all directions of momentum flow are present at once. The resultant "motion" is intricate consisting, at is does, of "rotations" about three distinct axes. Because of the partial analogue of the motion, on trying to pedal, that would be experienced by an initially stationary light cyclist in free space with three, relatively massive wheels mounted about three mutually perpendicular axes with each geared to simple multiples of each others phase (factors of 1 or 2) (also a twisting, turning, tumbling

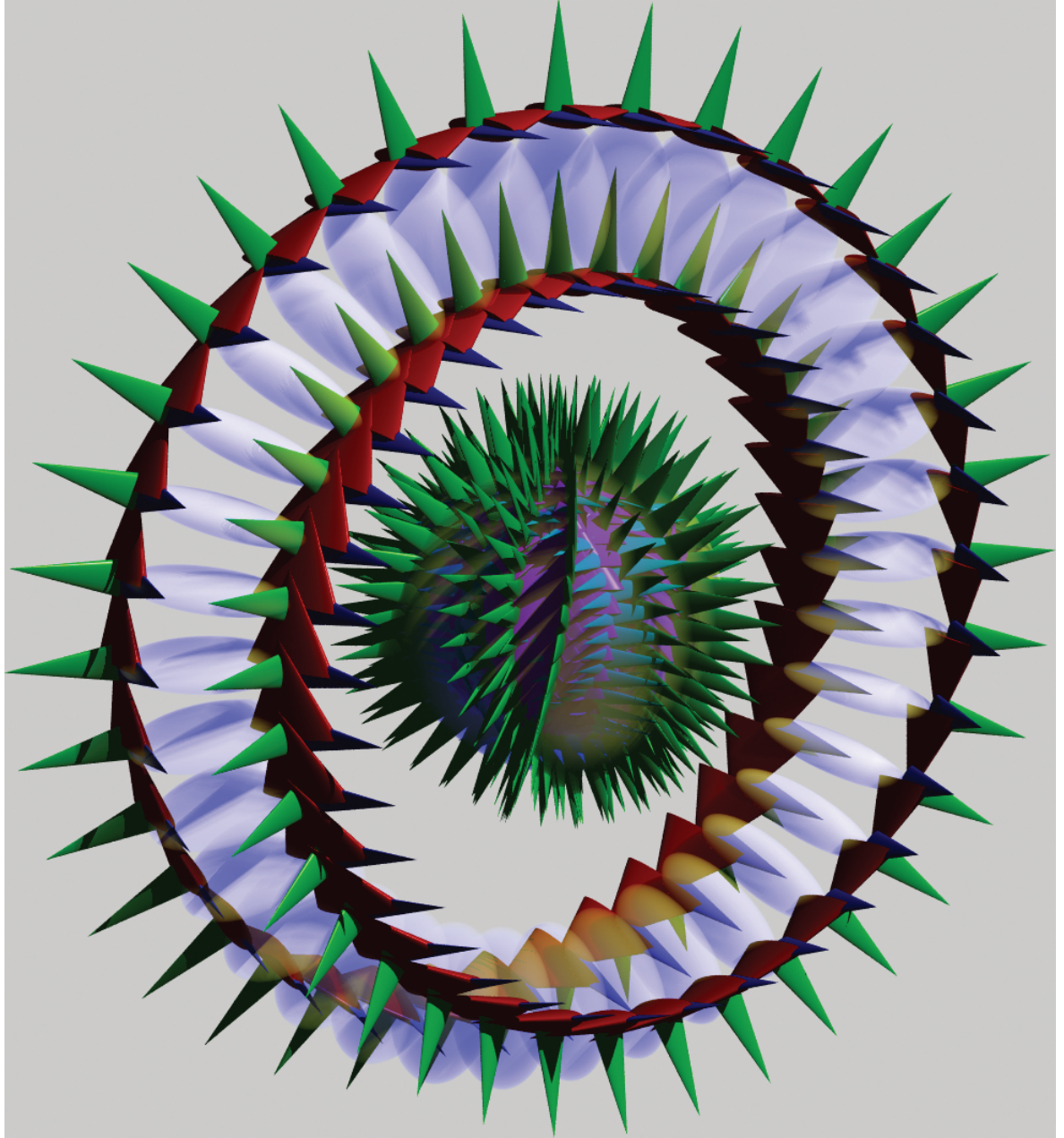


Figure 2. The four-dimensional quantum bicycle motion, illustrating the emergence of charge for a rest-massive particle. Expanded (twist and turn) and projected (twist, turn and tumble) view of the field distribution for the positron. The external double covering ring is an expanded view of the positron in Fig. 1 in field-momentum space and makes manifest the relationship to the internal toroidal topology in this space. The smoke disks represent the rest mass (scalar) component and correspond to two-dimensional slices through the hypersphere of the complete rotational motion in space-time. Each disk corresponds to a slice through the same sphere of rest-mass density (pivot) in normal space. The appearance in normal space is then spherically symmetric, as denoted in the central core. This is not shown to the same scale, and should correspond in size to inner toroidal radius (the smoke disk radius) which may be vanishingly small for the lowest energy configuration.

motion), Martin van der Mark and myself have dubbed the resultant four-dimensional inner motion of the elementary leptons the “quantum bicycle” (quicycle for short). If the rotational tumbling occurs at the Compton frequency, the magnetic field of one loop cancels the next, hence reducing the total energy of the configuration, dramatically. Such a resonant, harmonic motion once created would require a significant energy to unravel, leading to a stable, particle-like rest-massive charged object with internal spin. The complex motion would require an equal and opposite motion to conserve both momentum and the intrinsic internal angular momentum (the quantum spin). A cancellation of an electron vortex could be provided only by a positron vortex. In the absence of such a partner, the lightest such particles could not decay and conserve either charge, vorticity or angular momentum. The resultant is a coherent (in both of the main senses of the word) stable, self-sustaining and essentially spherical configuration of (root) mass and field: a paradigm for a quantum charged particle with intrinsic spin. The curvature defining the torus “size” is the ratio between the $P\vec{E}$ term and the $\vec{E} \times \vec{B}$ term in Eq. (16). Though this has a physical value of $\frac{4\pi}{\lambda_C} m^{-1}$, derived as the inverse radius of a double-loop of the whole mass-energy of the electron, it has been argued in earlier work[7] that this should be considered properly as remaining a bi-vector (in field-momentum space) and not a vector or scalar. The diagram at the central core is a result of the projection of these smoke disks onto the sphere of pivot at the core of the vortex-particle of the quantum bicycle motion. The configuration chosen in the illustration is that of the positron rather than the electron, since the outward-directed field is both easier to draw and to visualise. The average over all phases of the motion results externally in a spherically symmetric charged ball with an essentially zero resultant magnetic field and a radial electric field. There are four possible configurations of internal field as illustrated in Fig. 2. Two of these have outward-directed electric field, differing by the sense, left or right-handed, of the internal rotation of the twisting electromagnetic field about the internal momentum flow, the Poynting vector. These may be identified with the spin “up” and spin “down” positron. The other two have inward directed electric field, and differ only by a left or right-handed internal flow. These may be identified with the two spin states of the electron. The four minimum-energy configurations of the present model correspond to the four spinor basis in the set of solutions to the Dirac equation: spin up and spin down electron and positron. Note also that the internal frequency, being twice the Compton frequency, is identical to the zitterbewegung frequency in the Dirac model[4] and recently observed in experiment[19]. As mentioned in the introduction, a solution in Cartesian space such as Eq. (14) implies there exist similar solutions in any proper, conformal, orthonormal system of co-ordinates, modified only by well-known functions such as the half-integer Legendre polynomials[20]. The twisting, turning, tumbling wave with an internal 720-degree symmetry provides a paradigm for the internal structure of a fermionic charged quantum mechanical particle with internal structure. The configuration illustrated in Fig. 2 is a physical manifestation of a spinor wave-particle. That the relativistic wave-function of a pair of bosons, the photons described by Eq. 14, may fold into a pair of rest-massive charged fermions is the main result of this paper.

8. Discussion

So ubiquitous is the variation of the apparent electron size in experiment that one has become used to scarcely remarking on it at all. The author was lucky enough to work professionally with the European muon collaboration at CERN where leptons were observed to have a point-like interaction down to length scales smaller than $10^{-18}m$. A decade later the author was involved in the experimental measurement of the charge profile of individual electrons in the solid state where the charge of individual electrons was observed to be spread over length scales of tens of nano-metres consistent with the classic de Broglie form of quantum-mechanical wave-functions as they passed one at a time through a quantum point contact[21]. In atoms the electrons appear of an intermediate size of the order of Angstroms. Given this wealth of solid and apparently

contradictory experimental evidence where does this leave an estimate of the electron absolute “size”? For the removal of any further doubt about the extent of the theoretical problem: it is not consistent with experiment to contend that an electron should be a “point” particle in the sense of having (almost) zero extent: the electron cannot be a point and have finite spin, the electron cannot be a point charge and have finite mass-energy. To carry an intrinsic spin of $\hbar/2$ a simple calculation shows that the minimum electron size has to be at least $\frac{\lambda_C}{4\pi}$ even if all the mass-energy appears as pure light-speed momentum. Further, it cannot be smaller than the classical electron radius if the mass-energy in the electric field alone is not to exceed its total mass. The present work carries with it the possibility of understanding how an electron can appear so small in high-energy interactions, have a characteristic size consistent with its spin and yet expand remarkably in atomic and material systems. Firstly, one needs to understand that there is a world of difference between a point-like interaction of a particle and the notion of a point particle as an object of zero extent. What is observed in scattering experiments is that leptons display a scaling behaviour consistent with the scattering being independent of length-scale and hence point-like. This is not the same as the electron being a point particle. The key is to understand what the concept of “scaling” means. For a point-like interaction it is sufficient that the electron be spherically symmetric, have an inverse-square interaction and be a single, indivisible object rather than a collection of sub-objects held together by some unknown forces. All of these requirements are met by the model of the last section. To not resolve the size of an object, it is sufficient that the length-scale over which the field appears to extend is smaller than the effective size of any probe particle. It has been argued in earlier work that a self-confined photon displays a scale-free interaction with photons[10], and hence acts as a point-like object despite having an effectively finite size when at rest. The crucial argument is that, in electron-electron scattering, the apparent electron sizes scale relativistically in the same way as does the minimum wavelength exchange photon, yet remains smaller by a factor of at least 2π even in the most energetic head-on collisions. As the interaction energy rises, the effective apparent field dimension shrinks according to the same relativistic transformation as does the field of the photon probe, always remaining much smaller than the shortest possible wavelength exchange photon by the same ratio. As one attempts to hit it harder it appears simply ever smaller - its apparent size scaling with the centre of mass energy of the interaction according to the relativistic factor R defined above.

So much for the reason why the electron may appear “smaller” than it can possibly be in energetic elementary particle interactions. How about the reasons why it can then appear to be so large within atoms, molecules and the solid state? The details of precisely to what extent individual elementary particles expand depends on exact solutions of those systems and their inter-action with the environment and is in the realm of “further work” and a development to cover aspects of atomic physics. A mechanism, however, by which this may be possible at all may be outlined at the present stage of the development of the theory. Let us make the conjecture that the local pivot density can be reduced in electrons, atoms or material systems by a reduction in the absorption rate as will be discussed below in the context of the underlying nature of charge. In this situation the topological electron vortex can expand limitlessly while maintaining the same external field and quantum spin, both of which depend only on maintaining the fundamental inner toroidal topology[10]. This is what is envisaged to be the basic process of “elementary particle” size variation in atoms and in the solid state. In the charge-neutral hydrogen atom, the electron shields the proton and the proton the electron. In the composite electron-proton object there is little external field, so the interaction rate with the “rest of the universe” is dramatically reduced. Both electron and proton “blow up” dramatically as the relative internal pivot density (otherwise maintained by inter-actions) is reduced and as the external charge-field is reduced to (close to) zero in the (almost) perfectly spherical solution of the composite system that is the hydrogen atom. A simple calculation shows that the reduction

in total energy (the binding energy of Hydrogen) corresponds closely to the external electric field cancellation. The process continues further as one moves from atoms to molecules and to the solid state allowing the electron charge-spin-topology to be maintained and conserved, but over yet larger length scales. The vorticity, and hence charge as field divergence associated with the electron remains constant, the size does not. Indeed, in the solid state individual electrons are observed to become far “larger” than individual atoms[21]. This behaviour is part of the canon of (non relativistic) quantum mechanics, yet a plausible physical mechanism for the extraordinary flexibility of the size scale of “elementary” particles has not, until the advent of the present theory, been available, even though some of the aspects required have been discussed in general terms[22].

Another major mystery of physics is how fermions may arise from bosons and just what the fundamental underlying natures of fermions and bosons are anyway. Given the momentum of the constituent photon, it is straightforward to calculate the spin of the photon momentum in the double looped object[10]. This is just the characteristic length scale of the model $\frac{\lambda_C}{4\pi}$ times the initial photon momentum $\frac{h}{\lambda_C}$. That is, the numerical value of the spin is half integral: $\frac{h}{2}$. The total spin is then, by symmetry also half-integral since that of the initial photon rotates, averaging out to zero. By the spin-statistics theorem, therefore, the proposed object is a fermion. More fundamentally, as is clear from Fig. 2, the object is double-covering over the torus and returns to its starting configuration after a 720-degree rotation (rather than 360) in the space of an outside observer. That is, the object has the intrinsic internal spinorial symmetry of a fermion. The configuration illustrated in Fig. 2 is a physical spinor constituted by invariant root-rest-mass and (electromagnetic) fields. There are four and only four such solutions as the internal spin may twist leftwards (“up” say) or rightwards (“down” say) about the momentum flow and the resultant electric field may be inwards directed (electron) or outwards directed (positron) radially. The solution space is then identical that observed for the spin up and spin down electron and positron. In this picture, the fundamental difference between fermions and bosons is their inner topology, not just their rest-mass.

The Pauli exclusion principle is often couched in terms of identical fermions being “not allowed” to occupy the same quantum state or on requirements on the overall symmetry of the wave-function. Both of these sound more like the admonishments of a strict teacher than a proper physical law. A consideration of the internal interference of the fields in Fig. 2 on overlap leads to a physical origin of the force responsible for the exclusion principle and a limit to its strength as has been discussed in earlier work[18]. Briefly, it was argued there that imposing two parallel spin objects leads to an increase in total energy of the order of the particle masses - a huge repulsive force, whereas overlapping spin antiparallel leads to no increase at all. In the present context this argument needs to be updated as the invariant root-rest-mass must also be “shared” to give a zero-energy spin antiparallel configuration. This conjecture implies that there should be an upper limit of the exclusion energy potential of twice the total mass-energy of the system. This should be subject to experiment. There are hints already that this is at least of the correct order in experiments on high energy spin polarised scattering.[23, 24]

The underlying nature of charge and its quantisation is now discussed within the framework of the new model of the electron. A non-zero electric field divergence, a charge, can arise in the new model only if (self-sustaining) vorticity is introduced. The conservation of vorticity ensures the conservation of charge in the new theory. Conservation of charge and conservation of spin are linked, just as the magnitude of the quantised charge and the value of the quantised spin are linked, as calculated in earlier work[10]. Charged particles are produced in experiment only in vortex- anti-vortex pairs. If one’s current favourite theory does not provide a satisfactory description of the details of the dynamics of the process of the creation of photons from charged particles or vice-versa, then that theory is simply not complete in this respect. It should, by the scientific method, be either refined or replaced with something that does. Whatever replaces it,

however, should either lead to the replaced theories in some limit, or precisely replicate them in their realm of validity.

It is a simple matter to estimate the magnitude of the effective charge of a toroidal electromagnetic vortex from the magnitude of the trapped electromagnetic field and the photon momentum. The calculation relates two of the fundamental constants of nature, Planck's constant and the elementary charge, and allows either to be expressed in terms of the other. The comparison is based on a comparison of the outward -directed electric field in the vortex model, with that to be expected from a hypothetical "point charge" at the characteristic size of the model. The effective charge so calculated, viewed from outside the object proves independent of the size, mass and energy scales and hence fulfils at least one of the criteria needed for the existence of quantised charge. A comparison of the radial field for the case where the double-loops lie wholly on top of one another at $\lambda_C/4\pi$ provides one limit and this proves to be fractionally smaller than the elementary charge but within a few percent of it[10, 7]. If the internal mode structure is larger, the charge in the model would be increased, but only up to the order of a few tens of the elementary charge. It is encouraging that the calculated values span that of the elementary charge, and are of the right order of magnitude.

Depending on the conventional theoretical basis, charge has two aspects. Within classical field theory, it is the source of the electric field. As viewed from the standpoint of quantum electrodynamics, it is that thing which may emit or absorb mass-energy as photons between existing charged particles. A mechanism for the possible dynamics of this has been outlined in section 6.

The object depicted in Fig. 2 contains both dynamical invariant root-rest-mass and field. A rest-mass component and a field component. Within the theory, the field is confined by the mass, and the mass by the field. One can view the field component as being like a rapidly rotating light-speed sheepdog, corralling the rest-mass component which is its attractor and its centre, invariant under a Lorentz transformation, into a coherent whole. Such vortices provide stationary points, rest-massive frames which may emit or absorb photons in a process similar to that encapsulated in Eq. (14), thus effectively transferring rest-mass (pivot) between them. The emitter thus loses energy, increases in size, but retains its topology and external field. Similarly, the absorber gains energy and contracts. Such vortices are are pitchers and catchers of light according to Eqs. (13) and (14). An existing vortex may trap more field or emit part of it (subject, as usual, to energy-momentum conservation), in solutions corresponding to any energy as in Eq. (14). The process of photon absorption, for example, then merely "winds up" the internal vortex with an increase of energy frequency and a corresponding reduction in size, adding slightly to the internal field and pivot in equal measure. The charge, in field divergence terms, is mass-energy independent and depends only on the topology[10]. Charge is not then some fluid which an electron possesses, but a topological re-configuration of field due to a rest-mass component. Charge-charge interactions arise from photon exchange, an exchange of invariant mass-energy, intermediated by pure-field exchange photons, between existing such vortices. Provided only that the emission or absorption probability is, in dimensionless units, the fine structure constant. This is clearly a good basic starting point for the theory of quantum electro-dynamics. The new model then, fulfils both the aspects of charge discussed above. It is simultaneously a centre of radial electric field, and that object which may emit or absorb photons. So far so good but there remains a conceptual difference between charge as photon exchange and charge as field divergence. One might speculate that, just as inertial and gravitational mass are proportional but not necessarily equivalent, charge as field divergence, q_e and charge as quantised mass-energy exchange q_q are also proportional but not equivalent. One may write $q_e = a q_q$ where, for local charges in our reference frame the constant a is taken to be 1. Such a conjecture allows another possible interpretation for the nature of charge quantisation, as is now explored.

Consider the statement: “Hot electrons or systems interact more frequently than cool ones”. This does not seem very controversial. In fact it looks like a large part of the basis for the theory of thermodynamics. It seems to suggest, however, that “hot” systems are somehow “more” charged, not in terms of their external field which is size and hence energy scale independent, but in that they “pitch” light energy more frequently. In terms of the emission and absorption of photons, a higher energy system is higher frequency and emits more frequently, q_q is slightly larger, and q_e remains the same. The local quantisation of charge then has two aspects. In external field terms it arises from the resultant topology and is independent of size scale. In exchange terms it is a result of an equilibration between existing nearby particles and systems: “hotter” systems interact more frequently and hence tend to equilibrate locally. The effective “size” and hence mass of the electron would then depend on its external rate of inter-action, as discussed above in the context of the electron size scale in different systems. This would then mean that the constancy and “quantisation” of this aspect of charge then depends on the equilibration of local systems (charges) amongst each other, each centre equilibrating to the same effective interaction charge. This raises the possibility that the interaction aspect of charge q_q may vary by temperature, by epoch, or during creation before equilibration with local matter. This conjecture should be subject to experiment.

One further important feature of the earlier model was that it gave a physical explanation of the anomalous magnetic moment of the electron, which could be obtained from a proper consideration of the matching of rotating and non-rotating parts across the rotation horizon[10]. The value obtained, to first order, was the same as in quantum electrodynamics. That consideration carries over in the present model, as it has the same rotation horizon. Given that the charge, spin and anomalous magnetic moment of the double-looped object are close to those of the electron, it is considered self-evident that the four possible configurations of the simplest self-localised general vortices in Ξ_g should be identified with the states of the electron and positron. The body of physics, as it stands, is close to the truth in many practical respects. Any new theory should either encompass aspects of the old, improve on it, or provide a proper basis for its starting points. Like its sister model, the Dirac model, the new theory has four and only four rest-massive solutions. The spin up and spin down, electron and positron. Unlike the Dirac model, it explains both the underlying nature of quantised charge and the nature of half-integral spin. The four solutions, one of which is illustrated in Fig. 2, are physical spinors constructed of underlying root-mass and fields. These are the embodiment of a relativistic, charged wave-particle. Not in mass alone or field alone, but both. In summary the present first-order linear analytic theory encompasses the Maxwell equations, provides the physical basis for charge in quantum electrodynamics and is an extension in some respects of relativistic quantum mechanics.

9. Experimental tests

It may be considered that a theory extending classical electromagnetism and relativistic quantum mechanics and providing a basis for quantum-electrodynamics is reason enough that it should be considered to have some validity, but this is not the case: it must also agree with existing experiment, and should suggest new ones. The fact, however, that it fits seamlessly with existing theories means that new experiments cannot be proposed within the realm of either classical or quantum electrodynamics.

One striking new feature of the theory, the pivot, has not yet been isolated or observed in the laboratory. The new theory does not stand or fall on the observation of free pivot: it is perfectly possible that pivot exists only as a transitory state within existing material particles. The experimental isolation or observation of pivot, however, would constitute an extraordinary verification of certain aspects of its validity. At the same time there exists a deep mystery within current experimental science: the nature of “dark matter”. There are two ways that dark massive

material could hide: either it could interact so weakly that it simply passed through existing matter without being observed, or it could interact so strongly with matter that it was absorbed and integrated within any laboratory containing existing material particles. It is speculated that the latter may be the case, and that free pivot may constitute the so-far elusive “dark matter”.

Pivot, invariant root-mass-energy, is that quantity in the new theory that all quantum waves, with the sole exception of pure field waves, pass. Free pivot, in the theory, would be transparent to light, but strongly scavenged by existing material particles. Free invariant root-rest-mass, on encountering matter, would simply be absorbed, increasing its mass-energy, and hence leading to an increase in the internal frequency of the object. If there ever was any such stuff in the locality of the solar system, observations suggest that most of it has long ago been incorporated into existing matter. Such a process should manifest at low free pivot densities as an increase in low-energy heat of matter incident on free pivot. Clearly, locally in the solar system, the density of free pivot is vanishingly small. If such stuff existed in the past it has long ago been swept up by passing material. It may be possible, however, to generate free invariant root-rest-mass in the lab, for example by cancelling electromagnetic fields in the vacuum, where the invariant root-rest-mass component would then be formed in order to conserve energy as in Eq. (16). Such field cancellations should lead to excess di-gamma production in high energy high field regime. Some evidence that this may be the case has emerged in recent experiments at CERN[25]. At lower energies, one could attempt to probe a region where field cancellation is produced by high-power lasers with particle beams or low-energy photons. Plane polarised photons with a (half) wavelength similar to the region of invariant root-rest-mass, for example, should then be deflected in the electric-field plane, depending on the phase of the field.

10. Conclusions

A theory based on space, time and root-energy-density has been proposed. It introduces a sharpening of the principle of relativity, allowing only expressions, and solutions containing explicitly their proper space-time forms. The result proves to parallel classical electromagnetism, extend relativistic quantum mechanics and underpin quantum electrodynamics. The new paradigm allows relativistic photon wave-functions which scale smoothly from red-shifted oblivion to extreme blue-shifted gamma radiation, all governed by the single scale factor R . In the context of the new theory, the nature of charge and the topological relation between fermions and bosons has been investigated. Bringing in further a scalar root mass-energy term allows new kinds of root mass-field solutions. These are topological vortices in mass and field space. They are charged and possess the intrinsic properties of half-integer spin and are identified with the physical electron and positron. The new theory affords a hitherto unprecedented means of thinking inside the elementary particles of nature and expanding this thinking within a single unified theory of light and matter to include systems: atoms, molecules and solid state crystals. An example of this is the understanding of how and why electrons are observed to expand dramatically in atomic and material systems.

There is a possibility that root-mass may exist independent of material bodies. Such root-mass, though transparent to photons, would be strongly scavenged by existing material particles, incorporating it within themselves. This would manifest merely as an increase in energy. If no material particles were present, however, pivot could constitute a background mass-density in otherwise empty space. Such material, rest-massive yet transparent to massless photons, is a possible candidate for “dark matter”. It is not thought that the theory as outlined above is a full solution of Hilbert’s sixth problem. It is likely that the terms in spin, the tri-vector and the quadri-vector play as important a role as does the rest mass-energy term, the pivot, alone. Also it is considered unlikely that the conventional choices of ordering, handedness and sign adopted through history have been correct in every respect. For example, Hamilton and Maxwell used left-handed systems, as do developers of advanced computer simulations such as the program

POVRAY which has been used to generate the diagrams in this paper.

Acknowledgments

This work flows from a long and fruitful collaboration with my friend and co-pirate (eyepatch on the other eye) Martin van der Mark. I am grateful to Innes Anderson-Morrison for his work on the program ARPY, Phil Butler for his education in algebra and to Stephen Leary for his valuable contributions. I am also grateful to John Duffield, Tim Drysdale, Richard Gauthier, Richard De La Rue, Nick Bailey, Nick Green and John Weaver for useful discussions.

References

- [1] Leary S 2007 *Investigation of electromagnetism in a real Dirac algebra* Ph.D. thesis
- [2] Lounesto P 2001 *Clifford Algebras and Spinors* (Cambridge University Press) URL <https://doi.org/10.1017%2Fcbo9780511526022>
- [3] Doran C, Lasenby A, Gull S, Somaroo S and Challinor A 1996 Spacetime algebra and electron physics *Advances in Imaging and Electron Physics* (Elsevier) pp 271–386 URL <https://doi.org/10.1016%2Fs1076-5670%2808%2970158-7>
- [4] Hestenes D 2010 *Foundations of Physics* **40**
- [5] Williamson J G 2014 A new theory of light and matter *FFP14, Marseille, France*
- [6] Williamson J G and Leary S J 2015 Absolute relativity in classical electromagnetism: the quantisation of light *The Nature of Light: What are Photons? VI* ed Roychoudhuri C, Kracklauer A F and Raedt H D (SPIE) URL <https://doi.org/10.1117%2F12.2188363>
- [7] Williamson J G 2015 The nature of the photon and the electron *The Nature of Light: What are Photons? VI* ed Roychoudhuri C, Kracklauer A F and Raedt H D (SPIE) URL <https://doi.org/10.1117%2F12.2188259>
- [8] Jackson J D 1975 *Classical electrodynamics* (Wiley)
- [9] Gresnigt N 2007 *relativistic physics in the Clifford algebra $Cl(1,3)$* Ph.D. thesis
- [10] Williamson J G and Van der Mark M B 1997 *Ann. Fondation L. de Broglie* **22** 133
- [11] de Broglie L 1923 *Nature* **112** 540–540 URL <https://doi.org/10.1038%2F112540a0>
- [12] de Broglie L 1925 *Annales de Physique* **10** 22–128 URL <https://doi.org/10.1051%2Fanphys%2F192510030022>
- [13] Waite T 1995 *Physics Essays* **8** 6071
- [14] Hu Q H 2004 *Physics Essays* **17** 442–458 URL <https://doi.org/10.4006%2F1.3025708>
- [15] Gauthier R F 2007 FTL quantum models of the photon and the electron *AIP Conference Proceedings* (AIP) URL <https://doi.org/10.1063%2F1.2437555>
- [16] Robinson V 2011 *Particle Physics Insights* 1 URL <https://doi.org/10.4137%2Fppi.s7102>
- [17] Markou A 2012 *Physics Essays* **25** 547–552 URL <https://doi.org/10.4006%2F0836-1398-25.4.547>
- [18] Williamson J G 2012 Fermions from bosons and the origin of the exclusion principle *Proceedings of MENDEL 2012*
- [19] Gouanere M, Spighel M, Cue N, Gaillard M J, Genre R, Kirsch R, Poizat J C, Remillieux J, Catillon P and Roussel L 2005 *Ann. Fondation L. de Broglie* **30** 109114
- [20] Moon P and Spencer D E 1961 *Field theory for engineers* (van Nostrand)
- [21] Williamson J G, Timmering C E, Harmans C J P M, Harris J J and Foxon C T 1990 *Physical Review B* **42** 7675
- [22] Mead C A 2002 *Collective electrodynamics* (MIT press)
- [23] O’Fallon J R, Ratner L G, Schultz P F, Abe K, Fernow R C, Krisch A D, Mulera T A, Salthouse A J, Sandler B, Terwilliger K M, Crabb D G and Hansen P H 1977 *Phys. Rev. Lett.* **39**(12) 733–736 URL <https://link.aps.org/doi/10.1103/PhysRevLett.39.733>
- [24] Krisch A D 1979 *Scientific American* **240** 68–81 URL <https://doi.org/10.1038%2Fscientificamerican0579-68>
- [25] Atlas 2017 *Nature Physics* **13** 852858