# Absolute relativity in classical electromagnetism: the quantisation of light

J.G.Williamson[a]

[a] University of Glasgow, College of Science & Engineering, Glasgow G12 8LT, Scotland;

## ABSTRACT

A rigorous introduction of the underlying nature of space and time, through a sharpening of the principle of relativity, forces qualitatively new kinds of solutions in the classical theory of electromagnetism. A class of relativistic wave-functions are derived which are solutions to the first order, free-space Maxwell equation, These describe all photons from Radio to gamma waves and are governed by a single parameter: the exchange frequency, Though the theory remains that of classical, continuous electromagnetism, allowed travelling-wave solutions are quantised in that they come in “lumps” and are associated with a fixed angular momentum.

**Keywords:** light quantisation

## 1. INTRODUCTION

Since the early twentieth century theoretical effort has focused largely on the understanding of quantum mechanics and the development of gauge theories following on from the hugely successful theory of quantum electrodynamics. This paper picks up on an older path, that of classical electromagnetism and develops it within a relativistic mathematics designed to parallel experiment as closely as possible.The point of departure is represented by Maxwell’s classic text-book,[1] rather than more recent formulations of electromagnetism with a more complex superstructure.[2] The Maxwell theory has been re-cast in a minimal mathematics forced to parallel closely the experimental, relativistic, properties of space and time. The algebra generated is closely related to certain Dirac[3] and Clifford algebras,[4] but is more restrictive in some respects and somewhat generalised in others. In particular, a principle is adopted which forces all quantities, in all equations, to all orders to take their proper relativistic form. Because this extends and sharpens the principle of (special) relativity, this is denoted here the “principle of absolute relativity”. The severe constraints of this approach allow the usual four Maxwell equations to be written as a single equation (rather than in two pairs as in the more conventional approach[2]) in a form similar to that of the Dirac equation. The resultant equation may be expanded into a set of eight coupled differential equations, four of which take the form of the Maxwell equations. The resultant system is more general than that of either Maxwell or Dirac and has new kinds of solutions corresponding to both light and material particles. The latter are circulating, necessarily charged and with half-integral spin. Such solutions are beyond the scope of the present work. The aim here is to explore only light, leaving the investigation of the origin of the elementary charge and spin to a companion paper.

The solutions of the subset of the new theory corresponding to the free-space Maxwell equations alone will be explored. It is found that the only allowed propagating solutions are quantised. That is: the rigorous implementation of the principle of absolute relativity leads to the quantisation of light into photons. The structure of the paper is as follows. Firstly the Mathematical framework will be defined. Secondly a new set of equations, encompassing the Maxwell equations, will be derived. Thirdly, a new kind of wave function,incorporating the principle of absolute relativity, will be discussed. Fourthly, on the basis of this it will be shown that field only solutions correspond to the properties of light as observed experimentally. Electromagnetism remains continuous. Locally fields are unquantised, but propagation over longer distances forces solutions to be quantised as a consequence of the deeper principles of energy and field linearity and absolute relativity.

Further author information: (Send correspondence to J.G. Williamson)
J.G. Williamson: E-mail: john.williamson@glasgow.ac.uk, Telephone: +44 141 330 4923

## 2. OUTLINE OF THE THEORETICAL BASIS

Often, it is argued that a more general mathematics is more powerful than a simpler one. If one wishes to make an attempt to properly parallel reality, as in a solution of Hilbert's sixth problem for example, then one needs to find the simplest mathematics that parallels reality, just and no more. Here, an attempt will be made to keep the mathematics as restrictive as possible. This is not merely a philosophical choice: it is precisely the severe constraints imposed which lead to the necessity of travelling wave solutions of continuous classical electromagnetism being quantised in the following.

Most of the properties discussed here arise from the proper nature of space and time themselves. These properties are paralleled with a sub-algebra, isomorphic to the Dirac-Clifford algebra,[3–5] encompassing the relativistic properties of points, lines, planes, volumes and a hypervolume. This algebra has sixteen independent degrees of freedom. In addition a seventeeth degree of freedom is introduced, corresponding to a positive definite amount of energy. This energy may manifest in any of the geometric forms outlined above and may transform between them in a way well-described by the coupled differential equations to be derived.

A priori four (and only four) frame-independent unit elements are introduced. These elements represent unit lines in one dimension of time and three (orthonormal) dimensions of space. The elements themselves are frame-independent. Magnitudes or extents (which may transform between frames) are represented by positive-definite quantities with the appropriate units. These are used to express an amount of "stuff" (mass, energy or charge, for example) , or the apparent magnitude of (4- or multi-) vector elements in a particular frame of reference. So far this is five degrees of freedom, of which only one, the positive definite real quantities, represents a magnitude, the others being strictly unit 4-vector elements. Additional inner complexity arises rapidly under a proper,, physical, relativistic, definition of "multiplication" or "division" of the unit vector elements amongst themselves. Multiplication generates a unit point (for example of a unit line multiplied by itself), six unit planes (line times perpendicular line), four unit volumes and a unit hyper-volume, making, together with the four unit lines and the real magnitude, sixteen linearly-independent unit elements. The definitions are chosen such that these parallel (special) relativity precisely. For example a (unit) temporal line squares to the positive unit scalar element and a spatial line to the negative scalar. This sign is, strictly, an eighteenth degree of freedom. For practical purposes here, however, the sign may be subsumed into the reals, but it should be held in mind that it is, in fact, distinct. In summary: the mathematics is forced to parallel precisely that which is observed in experiment - not the other way round. The unit point, for example, is necessarily always a Lorentz invariant and is the"direction" of such things as the invariant mass. The four base vectors always transform as the components of a 4-vector for any quantity with this form. The 6 distinct unit areas transform, for example, as the six components of the electromagnetic field. Any other quantity with this form must transform in the same way. The (3 spatial) components of the tri-vectors transform as an angular momentum density. The quadri-vector is Lorentz invariant, but not invariant under inversion - and behaves in some respects as does the unit imaginary $i$, though it is quite distinct. In addition to these sixteen linearly independent"directions" a set of signs are required for each taking, potentially, the values $+$ and $-$ only. In particular cases, however, they not be required (or allowed) at all. For example, clearly, one needs to distinguish "forwards" and "backwards" in Cartesian space. It is at least debatable,however, whether a minimal description of reality will require both "forwards" and "backwards" in time. In the definition of the positive direction of the unit plane formed from the product of two perpendicular lines, one should distinguish the left-handed and right-handed choices with different signs. Note, in particular, that it has no meaning to add or subtract the unit elements themselves, but only the magnitudes which condition them: $1\alpha_0 + 1\alpha_0 = (1+1)\alpha_0 = 2\alpha_0$ (seconds, for example). The sign appearing in the addition or subtraction of real numbers is, again, different conceptually from the signs of the unit elements themselves. It is most important, to avoid confusion, to be mindful of the nature of the sign at hand. Also an extension of the simple basis here into the standard model may require more signs for various quantum numbers to distinguish such aspects as positive and negative charge, spin, lepton number and so on. It should be clear that the potential number of different "algebras" which may be defined in this way is rather large. Which, if any, is necessary as an element of the eventual solution of Hilbert's sixth problem is left to future work. The approach followed here has been to choose a system which works at the level of the Maxwell equations and which, further, corresponds as closely as possible with the conventions adopted in the standard textbooks.[2]

A four vector is written ($\mathbf{v} = a_0\alpha_0 + a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3$). Note carefullly that absolute relativity is imposed by using the $a_\mu$ to express a real magnitude (e.g. 3 Amps $m^{-2}$) or an extent (e.g. 42 metres) and the $\alpha_\mu$ to represent the proper unit-element vector form.

A product of these base unit elements with themselves is defined such that the unit time vector, $\alpha_0$ squares to the positive invariant scalar unity (the unit point) $\alpha_0^2 = \alpha_P$ and the three spatial vectors $\alpha_1$,$\alpha_2$ and $\alpha_3$ square to the negative scalar unity $\alpha_i^2 = -\alpha_P$. This is the point at which absolute relativity is introduced and is further all that is required such that all derived quantities transform correctly, relativistically, under all products and quotients The quantity $\alpha_P$ represents a physical point, not in size but rather as opposed to a line or a plane or a volume, in the algebra. Note that, for neither product, is the value assigned to the real number 1. The quantity $\alpha_P$ is distinguished, here, from the real or natural number unity (1) in that it is invariant under a Lorentz transformation and may take only the two values $\pm\alpha_P$. The positive value is idempotent such that $+\alpha_P^2 = +\alpha_P$. Here, the negative value is also taken to square to the positive unit scalar, $-\alpha_P^2 = +\alpha_P$. It is also perfectly possible, and more general, to work with the definition $-\alpha_P^2 = -\alpha_P$, but the definition adopted is that appropriate for real invariant masses which are always positive definite. It is worth noting that, properly, the multiplication (or division) of unit vectors, of magnitudes and of numbers are, in principle, three different kinds of operations. The first results in an object of a different form, the second in quantity with a different dimension and the third in merely a different magnitude. A consequence of the definition of multiplication above is that the square of a four vector is $(a_0^2 - a_1^2 - a_2^2 - a_3^2)\alpha_P$, a manifestly Lorentz invariant quantity, as it is experimentally. For the $\alpha_\mu$ taking the dimensions of a 4-momentum, for example, this is the positive-definite invariant mass.

The ordered product or quotient of one spatial unit element with another, for example $\alpha_1\alpha_2$ leads to a unit right-handed ordered spatial plane (bivector) element. This spatial plane is denoted $\alpha_1\alpha_2 = \alpha_{12}$. The reverse ordering gives a plane in the opposite (left-handed) direction, that is $\alpha_{12} = -\alpha_{21}$. There are three such right-handed objects: $\alpha_{12}, \alpha_{23}, \alpha_{31}$. Because this is a four-dimensional basis there are three further space-time planes, represented by products such as $\alpha_1\alpha_0 = \alpha_{10}$. Because of the properties of the base elements introduced above and the nature of the product, these elements transform relativistically as the magnetic ($\alpha_{ij}$) and electric ($\alpha_{i0}$) field elements which take this form in the following. This is a general and defining feature of the algebra being developed: anything with a particular unit element form inherits the relativistic transformation properties of that form. There are 4 tri-vectors representing unit volume elements ($\alpha_{123}$,$\alpha_{012}$, $\alpha_{023}$, $\alpha_{031}$). The latter three are a momentum density multiplied by a perpendicular unit vector, and therefore transform as the components of an angular momentum density. Finally, there is a quadri-vector ($\alpha_{0123}$) which, just as the scalar, is invariant under a Lorentz transformation but may change sign under other operations such as Hermitian conjugation.[6]

Several considerations should be noted. The system is non-commutative, hence the implicit ordering of quantities is important. In the sequel a system has been chosen which works, at least up to the derivation of the Maxwell equations. In principle, the elements derived from ordered multiplication or ordered division may be different. In particular, quantities of this form scale differently under a Lorentz transformation, as discussed below. The ordering of division (whether one divides by or divides into a quantity) introduces a sign change. Further, there are several choices to be made about the handedness and ordering of the operations between the various unit elements. In particular, the time element may be taken to come first or last (implying a change of sign and of handedness of the base elements in which it appears). Importantly, both choices give a same-handed set of products amongst each other ($\alpha_1\alpha_0 \times \alpha_2\alpha_0 = \alpha_0\alpha_1 \times \alpha_0\alpha_1 = \alpha_1\alpha_2$). This would imply that there was an intrinsic sign of and an intrinsic handedness between certain elements. The conventions adopted here work with the standard left-to-right ordering of products, the standard (right-) handedness of co-ordinate systems and the standard signs chosen for the directions of the electric and magnetic fields. This can equally be made to work with a left-handed basis. A comment is in order here: nature is intrinsically handed. The feeling of the author is that the left-handed choice is very likely to be more correct, though the right-handed choice has the advantage that it agrees with convention and hence is more comparable with results derived in most of the literature and in standard textbooks. Taking the convention that the base elements $\alpha_1, \alpha_2, \alpha_3$ are right-handed, this ordering, with space first then time, forms a right-handed set for angular-momentum products (such as $r \times p$), the reverse ordering a left-handed one. The conventional signs in the Maxwell equations then arise if one adopts the convention that the multiplication of a unit vector in the 1 direction into an inverse unit vector in the 2 direction has the reverse sign to the simple product. That is $\alpha_1/\alpha_2 = -\alpha_{12}$. It should be immediately apparent that, with these degrees of freedom, there is more than one way of choosing a consistent system at the level of the Maxwell equations. Further, conventionally, the scaling and sign properties are taken up by a real number factor (introducing positive and negative reals then) and, rather than introducing many more base elements than the 16 and obfuscating the simple development to follow, that approach will be followed here. Provided one is not working with addition or subtraction, but just with multiplication and division this is not an issue. Where it becomes an issue (in the addition of energies and fields), rather than being a problem it becomes a solution, as will be seen.

In the following, the proper form of quantities will be represented by a unit token with ordered lettering, thus $\alpha_{\mu\nu}$ represents a general bivector and, $\alpha_{0ij}$,$\alpha_{ij}$ and $\alpha_{i0}$ are right-handed tri-vectors, space-space bi-vectors and space-time bi-vectors respectively. Given this, the "direction" of a hypercomplex element is assigned to the way it transforms under a planar rotation.[6] For example, the unit volume $\alpha_{012}$ rotates in the same way as $\alpha_3$ hence, in cartesian co-ordinates it represents the "z" component of the angular momentum density. Here and in the sequel, Greek indices run from 0 to 3 and Roman from 1 to 3.

The multiplication and division of unit vector elements has been defined. The division of 4-vectors within the algebra is now discussed. Note firstly that the algebra developed is not a division algebra. There are many regions, apart from zero, where division is not defined. Primary amongst these is the vector itself. Consider the 4-vector case:

$$\mathbf{v} = \alpha_0 v_0 + \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \alpha_0 v_0 + \alpha_i \vec{v} \tag{1}$$

$$\mathbf{v}^{-1} = \mathbf{v}/\mathbf{v}^2 = v/(v_0^2 - v_1^2 - v_2^2 - v_3^2) = \frac{\mathbf{v}}{(v_0^2 - v_1^2 - v_2^2 - v_3^2)} \tag{2}$$

The inverse is in the same direction as the original vector, but with a different (real) scale factor, corresponding to the usual relativistic scaling. The over-arrow is used to denote the components corresponding to a conventional 3-vector. Here these are just three real numbers, with the proper (4-dimensional) unit elements being given by the $\alpha_i$ factors. Note, for the case of the space-time coordinates, the divisor corresponds to the invariant interval squared and that all inverses are scaled relativistically, by construction, according to this quantity. The underlying unit elements, when squared, give quantities of opposite sign, as they must relativistically. At the same time, if the real number factors for the magnitude of the spatial and temporal parts are equal, as they are everywhere on the lightcone for example, the interval goes to zero. Hence there is no inverse, not only at zero, but also in the crucial case of anywhere on the lightcone such that $(v_0^2 - v_1^2 - v_2^2 - v_3^2) = 0$. That is, the plane where division is undefined corresponds precisely to the physical limitations imposed by the speed of light. There are other combinations as well (such as that corresponding to the photon energy and momentum, for example) where division is undefined as well. Further discussion of where division is and is not defined is of great interest in itself, but not relevant to the simple cases discussed here. It is reserved for future work. Note that, for the case of the definition of the 4-vector derivative below, division is always well-defined. In that case the scaling is precisely unity in the frame of the derivative.

There is a feeling in some quarters that algebras which are not division algebras are somehow not well behaved.[5] This is true in the narrow sense of the properties of a "nice" mathematics. Here, it is precisely the proper scaling properties of 4-vectors in special relativity, perfectly paralleled in the present algebra, that leads to the new results. Far from being ill-behaved, this kind of behaviour is essential to precisely parallel the proper relativistic transformations of space and time, as observed in experiment, and to force the quantisation of allowed solutions.

For Cartesian co-ordinates a 4-vector 4-differential is defined within this framework as:

$$\begin{aligned} \mathcal{D} &= \frac{\partial}{\alpha_\mu \partial x_\mu} = \partial_\mu/\alpha_\mu \\ &= \alpha_0\partial_0 - \alpha_1\partial_1 - \alpha_2\partial_2 - \alpha_3\partial_3 = \alpha_0\partial_0 - \alpha_i\vec{\nabla} \end{aligned} \tag{3}$$

Note the imposition of the principle of absolute relativity by including the (quotient of) unit vector elements explicitly. It is this that leads to the change of sign in the spatial part above since $1/\alpha_0 = \alpha_0$ and $1/\alpha_i = -\alpha_i$. Note also that, though the differential is a special case of a division, the scale properties discussed above are simply unity, since the differential is taken with respect to each base unit vector element locally. The, emitter, absorber and the exchange field may each have different definition of "locality" and different scales for their local rulers and clocks, but the proper vector 4-differential works in each and every frame. If one were to define a single (scalar) differential with respect to an interval, this would be different. For any interval lying on the light-cone, for example, as the interval approached zero the result would tend towards infinity. The resultant does not tend to a finite quantity, but may be made "as large as you like" as one approached precise lightspeed. It may be speculated that this kind of behaviour may be part of the reason that zero-interval events dominate exchanges where the cross-section is vanishingly small, such as photon exchanges over intergalactic distances.

The 4-differential of a 4-vector potential yields field components. Writing a (proper absolute relativistic) vector potential as:

$$\mathbf{A} = \alpha_\mu A_\mu = \alpha_0 A_0 + \alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = \alpha_0 A_0 + \alpha_i \vec{A} \tag{4}$$

It should be noted that there is a possibility, in the full theory to be developed below, that the 4-trivector may also yield field components.The 16 $(= 1 + 3 + 3 \cdot 2 + 3 \cdot 2)$ terms of the 4-derivative of the 4-potential $\mathcal{D}A$ may be gathered together and written as:

$$\mathcal{D}\mathbf{A} = \alpha_P(\partial_0 A_0 \quad + \quad \vec{\nabla} \cdot \vec{A}) - \alpha_{i0}(\partial_0 \vec{A} + \vec{\nabla} A_0) - \alpha_{ij} \vec{\nabla} \times \vec{A} = P\alpha_P + F\alpha_{\mu\nu} \tag{5}$$

which is the sum of a scalar (pivot) part $P\alpha_P$ and a bivector (field) part $F\alpha_{\mu\nu}$.

In Eq. (5) the term in $\alpha_{i0}$ is usually identified with the electric field $\vec{E} = -\partial_0 \vec{A} - \vec{\nabla} A_0$ and that in $\alpha_{ij}$ with the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$. Some care is required with the signs in relating these quantities to the conventional electric and magnetic fields. Taking the convention in Jackson[2] $F\alpha_{\mu\nu} = E_i\alpha_{i0} - B_i\alpha_{jk}$). The standard electric field then maps to the set of three ordered right-handed space-time unit elements $\alpha_{10}, \alpha_{20}, \alpha_{30}$ and the magnetic field to the terms $\alpha_{23}, \alpha_{31}, \alpha_{12}$ respectively.

Over each of the sixteen multivector-quantities defined above, a general dynamical multi-vector field $G$ is defined over a scalar term $P$, a vector potential $A$ a field term $F\alpha_{\mu\nu} = E_i\alpha_{i0} - B_i\alpha_{jk}$, a trivector potential term $T$ and an eventual quadri-vector potential $Q$ such that: $G = P\alpha_P + A_0\alpha_0 + A_i\alpha_i + E_i\alpha_{i0} - B_i\alpha_{jk} + T_k\alpha_{0ij} + T_0\alpha_{123} + Q\alpha_{0123}$. In an obvious notation, the constant terms are defined as $C = C_P\alpha_P + C_0\alpha_0 + C_i\alpha_i + C_{i0}\alpha_{i0} - C_{jk}\alpha_{jk} + C_{0ij}\alpha_{0ij} + C_{123}\alpha_{123} + C_Q\alpha_{0123}$.

Writing, by analogy with the form of the Maxwell equation $\mathcal{D}F = J$, $\mathcal{D}G = C$, and again using the conventional 3-space patterns for reference, one obtains from the odd terms a generalisation of the Maxwell equations as:

$$\begin{aligned}
\alpha_0(\vec{\nabla} \cdot \vec{E} + \partial_0 P) &= C_0\alpha_0 && (6)\\
\alpha_{123}(\vec{\nabla} \cdot \vec{B} + \partial_0 Q) &= C_{123}\alpha_{123} && (7)\\
\alpha_i\left(\vec{\nabla} \times \vec{B} - \partial_0 \vec{E} - \vec{\nabla} P\right) &= C_i\alpha_i && (8)\\
\alpha_{0ij}(\vec{\nabla} \times \vec{E} + \partial_0 \vec{B} + \vec{\nabla} Q) &= C_{0ij}\alpha_{0ij} && (9)
\end{aligned}$$

and four further equations in the even terms:

$$\begin{aligned}
\alpha_P(\vec{\nabla} \cdot \vec{A} + \partial_0 A_0) &= C_P\alpha_P && (10)\\
\alpha_{0123}(\vec{\nabla} \cdot \vec{T} + \partial_0 T_0) &= C_Q\alpha_{0123} && (11)\\
\alpha_{i0}\left(\partial_0 \vec{A} + \vec{\nabla} A_0 + \vec{\nabla} \times \vec{T}\right) &= C_{i0}\alpha_{i0} && (12)\\
\alpha_{jk}\left(\partial_0 \vec{T} + \vec{\nabla} T_0 - \vec{\nabla} \times \vec{A}\right) &= C_{jk}\alpha_{jk} && (13)
\end{aligned}$$

In the first set of four equations corresponding to the Maxwell equations, the main new feature is the introduction of two new dynamical terms $P$ and $Q$ transforming (under a Lorentz transformation) as invariant masses. The constant terms on the right correspond to the electric and magnetic charge and current. Setting $P$

and $Q$ zero, but setting $C_0$ to the charge and $C_i$ to the current density these reduce to the standard, inhomogenous Maxwell equations. Note that, in this limit, all four Maxwell equations are present at once in the present formalism, with all the correct signs, in contrast to the usual derivation not using the principle of absolute relativity.[2]

In the second set of four equations the main additions are the presence of the tri-vector terms $T$. These represent the possibility of introducing a 4-tri-vector potential as well as the conventional 4-vector potential into an extended theory of electromagnetism. This may be expected to be of value in understanding the underlying nature of angular momentum in particles. Further these equations express (potential) degrees of freedom. With $C_P = 0$, equation (10) is just the Lorenz gauge condition and one obtains other conventional gauges by setting this constant to other quantities. In other words, $C_P$ non-zero expresses a gauge degree of freedom. Here, there are other constants which, if expressed, would introduce new physics. In particular, the dual term $C_Q$ expresses a further degree of freedom. In principle, the odd set and the even set should both constrain the physics, but the even set will not be used in deriving the main results of this paper. Indeed, the main results here will be obtained by demanding that solutions satisfy the standard set of free-space Maxwell equations alone ($\mathcal{D}F = 0$). The full set will be used in developing the linear, first order wave-functions used to construct the solutions. Note that, with $T = 0$ the final two equations are just the standard expression for the electric and magnetic field in terms of the vector potential.

Within the formalism, the physical effect of the new term $P$ is to allow a curvature of the momentum transport direction. If non-zero, this leads to the possibilities of a pivoting of the field flow around the mass leading to new kinds of self-confined circulating solutions with rest-mass.[9,13] These solutions may underpin the underlying quantised nature of charge.[10] The possibility that this new framework constitutes a new general, linear theory of light and matter, treating leptons and photons on the same footing, will be explored further in a companion paper.[14] Here, all the constant terms on the left will be set to zero. This corresponds to the free-space (Lorenz gauge) condition for the conventional Maxwell equations alone. This is the appropriate framework for the description of the photon.

## 3. A NEW PHOTON WAVE-FUNCTION

Using the algebra new kinds of wave function may be generated with properties more strongly constrained than is possible conventionally. There are a plethora of such solutions. Complexity has arisen rapidly from the simplicity of the four-dimensional basis due to the properties ascribed to multiplication and division. As a prequel, it is worth exploring these new solutions in general.

Conventionally, one often writes wave functions introducing a complex scalar $i$ and exploiting the property that:

$$e^{i\theta} = e^{\theta i} = (e^{i})^{\theta} = (e^{\theta})^{i} = cos(\theta) + isin(\theta) \tag{14}$$

The ordering and nesting of the exponents in the equation above is unimportant for complex numbers, as all factors commute, but will prove crucial in the more complex discussion to follow. As is well-known, a non-relativistic wave function propagating in the z direction may be written:

$$F_{NR} = Ae^{i(kz-\omega t)} = Ae^{i\theta} \tag{15}$$

Where $k$ is the spatial frequency (wavenumber in z) and $\omega$ is the temporal angular frequency. Such forms have wide practical application. They may be further modified by well-known generalisations of the harmonic functions, Bessel functions, spherical harmonics, half-integral Legendre polynomials and so on to describe exponential-like solutions in cylindrical, spherical and toroidal systems. Despite their power and elegance, they have one major flaw if one wishes to use them in a relativistic theory: space and time appear in the combination $(kz - \omega t)$ as a (Lorentz) scalar factor. Since space and time, however, transform differently under a general Lorentz transformation, such a wave function does not reflect the differences between space and time properly. Such wave-functions do not, therefore, conform to the principle of absolute relativity. Rectifying this by imposing the proper relativistic transformations of space and time in absolute relativity at all levels, most especially in the exponent, leads to qualitatively different kinds of solutions. It is the rigour of forcing all instances of space or time to have their proper form, wherever they occur, that leads to allowed solutions of the Maxwell equations corresponding more closely to the physical photon, in that light must come in "lumps" that scale in total energy with the frequency. This section will develop the general case, the next- the photon.

In the present formalism, there is no single simple complex scalar $i$, but several quantities which may play the same role in describing travelling wave solutions. Three of the base unit elements and seven of the unit elements derived from these square to negative unity (explicitly these are: $\alpha_1, \alpha_2, \alpha_3, \alpha_{12}, \alpha_{23}, \alpha_{31}, \alpha_{012}, \alpha_{023}, \alpha_{031}, \alpha_{0123}$). Any of these may be used to describe travelling waves. For example, by analogy with complex numbers one may expand an exponential with $\alpha_{12}$, corresponding to a rotation of angle $\theta$ in the 12 plane as:

$$e^{\alpha_{12}\theta} = \alpha_P cos(\theta) + \alpha_{12} sin(\theta) \tag{16}$$

In the physical association made above, this would describe an oscillation back and forth between a rest mass component ($\alpha_P$) and a magnetic field component($\alpha_{12}$). Such a formalism is descriptive in a similar way to complex numbers. Using the scalar $\alpha_P$ and $\alpha_{0123}$ alone provides an even more precise parallel, since the sub-algebra containing this pair alone is isomorphic to complex numbers. Such exponents will be denoted in general as hypercomplex exponents. Though this may sound like some progress, eq. (16) retains the problem of general covariance alluded to above and such solutions are not necessarily proposed as representing a physical process (governed by a 4-vector derivative) as the proper elements corresponding to neither space nor time are present. To describe physics, they would require a bi-vector or quadri-vector derivative in an angular measure to operate. This would retain the essential feature of the deficiency sketched above, that space and time are treated identically. They serve merely to point the way to further progress.

Note that $\alpha_P$ has been used in the expansion above. This is because energy conservation considerations require that any physical wave should transform between elements of substance constituting, at the very least, an equal integrated energy. It would be preferable, of course, to require a local microscopic conservation of energy (and momentum), if this could be defined. In either event, this means that both terms should square to an energy density rather than to a simple number. This is one reason why at least two distinct kinds of scalar unity must be considered. The unity $\alpha_P$ appearing in hypercomplex exponentials describing unitary transformations in nature is necessarily different to the unit real number 1. Moreover, it will appear that consideration of the non-commutativity of elements may lead to a kind of black-body quantisation, as will be discussed in the following.

To make proper progress, a second extension is required such that elements may be nested with each other leading to a combined motion observed as a wave. This leads to a far richer structure than is available in a merely complex algebra. The new axiom requires the inclusion of the proper (in the Lorentz sense) relative transformation properties of space and time directly into the exponential. This may be achieved by associating the proper unit element directly with the appropriate propagation direction in the hypercomplex exponent, as is now shown. For example, for propagation in the 3 (z ) direction, solutions are sought for some appropriate unit element, denoted $\alpha_?$, of the form:

$$F_{SR} = Ae^{(\alpha_3 kz - \alpha_0 \omega t)\alpha_?} \tag{17}$$

Absolute relativity is imposed in that space and time appear again with their proper form (here: $\alpha_3$ and $\alpha_0$). By analogy with the expansion of real exponentials to complex exponentials, an element $\alpha_?$ is required to convert this to a travelling wave-function. Note that though equation (17) is a conceptual extension, it also embodies a physical restriction in that the factors corresponding to space and time are forced to have their proper relative form.

For a wave-like overall solution $\alpha_?$ is required to be some unit element which ensures that both the spatial and the temporal element of the development is governed by a unit element squaring to negative unity. Within the principle of absolute relativity it is axiomatic that space and time, and any other quantities such as angular momenta, should appear everywhere with their proper form.

By inspection, it is apparent that substituting neither $\alpha_P$, nor the real number 1 for $\alpha_?$ leads to a travelling wave solution. In both cases the temporal development will square to positive unity, leading to falling exponential-like solutions rather than waves. The question is: by analogy with the extension from ordinary to complex exponentiation outlined in equation (14), is it possible to pre- or -post multiply the exponent by some other unit element, corresponding to exponentiation by or exponentiation into the whole expression? For a given propagation direction there are six unit elements $\alpha_?$ which achieve this, each leading to new kinds of wave-particle solutions. For the particular case of the 3 direction in eq. (17), these are explicitly: $(\alpha_{012}, \alpha_{23}, \alpha_{31}, \alpha_{123}, \alpha_{10}, \alpha_{20})$. The first three themselves square to negative unity; the second three to positive unity. This means the first three afford the possibility of inserting a scalar phase factor, in harmony with the multi-vector component, into the hypercomplex exponent. Of these, only $\alpha_{012}$ (corresponding to the introduction of a unit vector in the direction of the angular momentum) leads to the possibility of substituting for $A$ in equation (17) above, a pure field solution. This is presented below. The corollary to the principle of absolute relativity here is then that travelling wave solutions in space and time require a unit angular momentum in order to propagate. Such solutions are then a first order solution of the Maxwell equations - describing all six components of the electromagnetic field in any proper frame - but they are necessarily associated with a unit angular momentum.

Equation (17), with $A$ scalar, though it represents a wave, is not itself of solution of the free-space Maxwell equations $\mathcal{D}F = 0$ as it contains terms transforming as a rest-mass as well as field terms. It is, however, a solution of the more general set of dynamical equations with the constant terms zero. That is: $\mathcal{D}F_{SR} = 0$, as is readily verified by substitution. It is tempting to associate such a wave-function directly with a massive source particle such as an electron. This is partially true, but this form is still too simple to fully encompass the complexity of such particles as the scalar pre-factor is far too simple.[10,13] Further developments along these lines will be dealt with in the companion paper.

It is worth noting in passing that, of the other five, the elements $\alpha_{23}$ and $\alpha_{31}$ also lead to wave solutions of the general equation dG=0, and these may correspond more closely to elements of electron-like and positron-like solutions. The remaining three possibilities may also be associated with light-like and particle-like solutions and may indeed be the primary- initial or lightest- solutions. Curiously, the dual bivector pair ($\alpha_{10}$ and $\alpha_{20}$) do not lead by themselves to a magnetic monopole-like but also to an electric monopole-like fields, as can readily be verified by substitution and expansion. Magnetic monopoles may be described, however, by introducing more complicated terms involving a product with the pre-factor $A$. The $\alpha_{123}$ case may be associated with a precursor to the electron-positron pair in the creation process as it resembles a twisted-mode solution, the solution obtained by overlapping counter-propagating right-right or left-left circularly polarised light. This is the configuration for the creation of a particle-antiparticle state at spin zero at sufficiently high energy. It is also possible (by choosing an appropriate pre-factor and/or the relative propagation direction of space and time) to associate these with the primary photon-like and electron-like solutions. This opens up the possibility that the other set, which may introduce mass through the scalar term, may be involved in the description of the weak interaction. Again, the development of these speculations will be left to future work.

In popular expositions of relativity one often talks of rulers and clocks, these often being held by idealised "observers". In reality, there are no observers, only the emitting and absorbing particles or systems and the intermediating photons themselves. Confusion is often introduced in arguments by ascribing "knowledge' to n external observer that it could not possible have.

Consider an idealised system in three frames: an emitter, an absorber and an intermediating photon. Each frame has its own scale of space, time and frequency. Its own "rulers" and "clocks" and its own scale of energy or, equivalently, frequency. In a general "event" where a photon is exchanged between two particles, the particles may be in very different Lorentz frames. These frames, and hence their scales, will also change due to the effect of the exchange. Let the scale-change be denoted by $R$. For example, a photon in one frame may have a particular energy, frequency and wavelength. In another (blue shifted) frame where the energy (and hence the frequency) increases by $R$, the wavelength decreases by $1/R$. It is enlightening to write this scale-factor $R$ in terms of the usual relativistic $\beta$ and $\gamma$ factors:

$$R = \sqrt{\frac{1+\beta}{1-\beta}} = \gamma(1+\beta), \quad 1/R = \sqrt{\frac{1-\beta}{1+\beta}} = \gamma(1-\beta), \quad \omega' = \omega R, \quad \lambda' = \lambda/R \tag{18}$$

With:

$$\beta = v/c = \frac{R^2-1}{R^2+1}, \quad \gamma = \frac{1}{\sqrt{1-v^2/c^2}} = \frac{1}{2}(R+\frac{1}{R}) \tag{19}$$

Note that the last relation above means that the gamma factor is the average of the increase in energy of the light travelling against the motion, with that travelling with the motion. One may conclude that the Lorentz scaling of the mass of material particles is just that of the energy of light in a box. Further, 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 $|E| = |B|$ , they transform relativistically as:

$$E' = \gamma(E+\beta B) = RE(=RB), B' = \gamma(B+\beta E) = RB(=RE) \tag{20}$$

That is, for light, the fields transform in the same way as does the frequency and energy: linearly with $R$.

The formalism to write down a new, fully relativistic solution to the first-order Maxwell equation in free space ($\mathcal{D}F = 0$) is now complete. For the simple case of a propagating free-space electromagnetic wave, forcing $z$ and $t$ to take the proper form $\alpha_3$ and $\alpha_0$ respectively, a single photon solution of a left circularly polarised electromagnetic wave, travelling in the the $+z$-direction and transmitting a quantum of energy $\mathcal{E}$ in the centre of mass frame may be written:

$$F_L = H_0 U_F R\mathcal{E}(\alpha_{10}+\alpha_{31})e^{\frac{\mathcal{E}}{\hbar}R(\alpha_3\frac{z}{c}-\alpha_0 t)\alpha_{012}} = F_0 R(\alpha_{10}+\alpha_{31})e^{R(k\alpha_3 z-\omega\alpha_0 t)\alpha_{012}} = \mathcal{F}\mathcal{W} \tag{21}$$

This has a pre-factor part representing the initial (or final) field configuration $\mathcal{F}$, and a hypercomplex exponential wave-function $\mathcal{W}$.

The real-number constants $c$ are the (scalar) speed of light and $\mathcal{E}(=\hbar ck=\hbar\omega)$ the (scalar) quantum of energy transmitted in the centre-of-momentum frame respectively. $U_F$ is a universal constant, taking the same value for all photons, converting to field units. $H_0$ is a distribution function representing the spread of field or energy over phase, whose square integrates to unity. This represents the number of cycles in phase over which , for the photon, the emitter and absorber remain in mutual phase coherence. This is an invariant and is the same in all frames, right up to the limit of light-speed where the integrated energy goes to zero. The single parameter R is that factor which determines the scales of energy, frequency, length and time for the same photon in any Lorentz frame. Taking as a reference that R is unity in the centre of momentum frame then $\mathcal{E}$ determines the proper magnitude of the energy-momentum transmitted. The factor $\mathcal{E}R$ is then the energy in each relevant frame. The factor of $R$ in the exponent pertains to the relativistic transformation of "rulers" and "clocks". The proper reference "ruler" and "clock" for any given photon exchange event scales with the centre of momentum energy. This is just the wavelength and frequency of the photon in the proper (centre of momentum) frame. The factor of $R$ in the pre-factor corresponds to the proper relativistic transformation of fields in the emitter and absorber frames. Both field and frequency scale linearly with $R$. In the centre of momentum frame the proper frequency is the energy divided by Plank's constant $\omega_0 = \frac{\mathcal{E}}{\hbar}$. Note that the energy density in the field is proportional to field squared. Explicitly, $D(\mathcal{E}) = \frac{1}{2}\epsilon_0(E^2+c^2B^2)$. This means the wave-function in equation (21) may be converted to square-root energy density units by the simple expedient of defining another universal factor $U_{\mathcal{E}}$ in place of $U_F$. This, then, is a fully relativistic wave-function giving the energy "probability density" in any desired frame. This raises the question of why the factor for energy appearing in the pre-factor should be $\mathcal{E}$ and not $\sqrt{\mathcal{E}}$. The reason is that, experimentally, both energy and field must add linearly. If one accepts this primary experimental fact, then it is space and time themselves which must deform, relativistically, to accommodate these deeper principles of linearity. That is exactly what the Lorentz transformation achieves, as is sketched above. It is exactly that transformation which linearises the addition of both energy and field in the allowed solutions. If the scaling factor is 2, for example, the relativistic transformation is such that the field doubles, the energy density quadruples, but the length of the wave train in the new frame halves (due to the observed deformation of space relativistically), giving a linear increase in the total energy with frequency overall, as is observed experimentally. It is the number of cycles of phase which remains the same in all frames-not the length or time. It is the stringent constraints of linearity of energy and of field, together with this, that forces allowed solutions to be only those where the frequency in the exponent corresponds to the proper magnitude of the energy in the pre-factor. One can fill in any value of energy from radio-waves to high energy gamma photons to the wave-function above - but this must affect both exponent (frequency) and pre-factor (field) proportionately.

The Wave-function $\mathcal{W}$ alone is not, strictly, a solution of the conventional Maxwell equations, as it contains terms transforming as an invariant mass as well as terms transforming as fields. It is, however, a solution of the more general set of equations discussed above such that $\mathcal{D}\mathcal{W} = 0$. For the pure-field cases of many wavelengths observed to exist physically, in the product $F_L = \mathcal{F}\mathcal{W}$, the mass terms cancel leaving pure field alone, so that the whole expression is then a solution of the free-space Maxwell equations $\mathcal{D}F_L = 0$. What this means, physically, is that, $\mathcal{W}$ may be an element of the proper relativistic wave-function of the (rest-massive) emitter and absorber. If so, such physical functions may combine with initial and final fields such that they effectively propagate a field packet of arbritary total energy. Such fields are propagated at light-speed only if the initial (emitter) and final (absorber) fields have equal and perpendicular magnetic field components and are such that the total energy in their frames (note carefully that these are usually different) is proportional to the local wave-function frequency $\nu_l$ (where the suffix $l$ denotes the local frame under consideration). Denoting emitter $e$ absorber $a$ and photon $p$ one needs to consider the concept of locality for the same photon in three frames. This is a coherent tri-locality. The locality of the emitter, the locality of the absorber and the locality of the exchanged photon.The photon emerges from the emitter with energy $\mathcal{E}_e = h\nu_e$. For the absorber the same photon arrives with energy $\mathcal{E}_a = h\nu_a$. For the co-moving (at least nearly lightspeed) frame of the photon, the energy $\mathcal{E}_p = h\nu_p$ tends to zero, the wavelength tends to infinity, but the transformations of absolute relativity place each phase value in $H_0$ at (very nearly) the same point in space-time for both the emitter and absorber. Locality, for the photon, may span vast tracts of space and time. The absorber "sees" a packet of positive energy arriving from its past. The emitter "sees" a packet of negative energy leaving into its future. This is the same photon. Positive energy backwards in time is the same thing as negative energy forwards in time. In any event both happen. Causality is in the direction of the energy transmitted. From emitter to absorber. That fixes the "arrow of time" for each individual event. Taken together the picture is of a resonant, coherent, smooth exchange of a packet of mass energy characterised by the frequency alone. This is exactly what is observed experimentally.

Elements of the application of the principle of absolute relativity have appeared seven times in constraining the form of equation (21). Space has been inserted with its proper form ($z$ here) in the direction of propagation direction as $\alpha_3$. Time appears associated with $\alpha_0$. The proper "direction" of the angular momentum around the propagation direction $\alpha_3$ is $\alpha_{012}$. The electromagnetic field must obey strict constraints in order for equation (21) to be a solution: the starting fields must be perpendicular, $E_x$ and $B_y$ here in proper "directions" $\alpha_{10}$ and $\alpha_{31}$ and they must be of equal magnitude; otherwise the rest-mass terms do not cancel to leave a zero rest-mass pure field capable of light-speed propagation. The (scalar) energy must then appear twice to be consistent with the relativistic laws: once in the pre-factor (expressing the linearity of field under a Lorentz transformation with $R$) and once in the exponent (expressing the linearity of energy and its proper transformation relativistically). If all of these conditions are satisfied, the resulting propagating part is necessarily quantised, according to equation (21).

Equation (21) is a universal, relativistic, wave-function describing photons of any energy. Varying $\mathcal{E}$ from zero to infinity one obtains any photon of any energy, all with the. same angular momentum. Conversely, the same photon, viewed from different frames characterised by relativistic scaling factor $R$, will appear to have different frequencies and energy (right down to the zero-energy limit) but the same angular momentum. One can , as a photon emitter or absorber, be in any arbitrarily blue or red-shifted frame, the wave-function takes the same form, differing only by the change in the relative scale of frequency, energy and length as described by the relativistic scaling factor $R$. In every frame however, these properly relativistic wave-functions are necessarily quantised with a characteristic frequency $\nu_l$ and an integral energy given by $\mathcal{E}_l = h\nu_l$. The new, fully relativistic, quantised wave-function equation (21) is the main result of this paper.

Usually one looks for solutions of the second order equations eliminating one of the fields in favour of the other. equation (21) is a solution of the first order equation $\mathcal{D}F = 0$ directly. It describes both electric and magnetic fields simultaneously. The development of these in time and space parallel the field transformations expressed by the Maxwell equations. Under a Lorentz transformation the whole solution scales in energy proportional to the frequency alone, as is observed in experiment. Taking the frequency to zero, the energy also goes to zero. Because of the construction, all elements, in any frame and to any order of expansion, scale properly relativistically. This is a result of the rigorous implementation of the principle of absolute relativity, particularly in that the proper relative form of space, time and angular momentum appears in the exponent and that the energy appears both in the exponent and in the pre-factor.

It is also worth noting that, in contrast to conventional wave-functions of the form of equation (15), the development of the wave-function in space and time is different. In space one has a wave, in time one has a rotation. As one moves in space there is an alternation between electric and magnetic field - just as described by the first order Maxwell equations. Sitting at one point in space, and allowing time to pass, one has a rotation of the field vectors. The new wave-functions, in and of themselves, parallel the physical development of the field components in the Maxwell equations far more closely than do conventional solutions.

To make the connection with the form encountered in standard textbooks,[2] equation (21) may be readily be expanded in any particular frame. For the conditions corresponding to experimentally observed photons, the non field (scalar and quadri-vector) terms in the exponential part cancel. Setting $F_1 = H_0 UR\mathcal{E}$ and $ck = \omega = \frac{\mathcal{E}}{\hbar}$ one obtains:

$$F_L = F_1[(\alpha_{10} + \alpha_{31})\cos(kz - \omega t) + (\alpha_{23} - \alpha_{20})\sin(kz - \omega t)] \tag{22}$$

This describes electric ($\alpha_{i0}$) and magnetic ($\alpha_{ij}$) fields rotating in time in a plane perpendicular to the direction of momentum transport and transforming in space from magnetic to electric and vice-versa. The resultant field configuration is that shown in Fig. 1. It is identical that found in any elementary textbook on electromagnetism for a left-handed circularly polarised wave. This is comforting: provided the constraints are satisfied, despite its apparent complexity, the new wave-function simply expands to that form measured in experiments and familiar from elementary text-books.

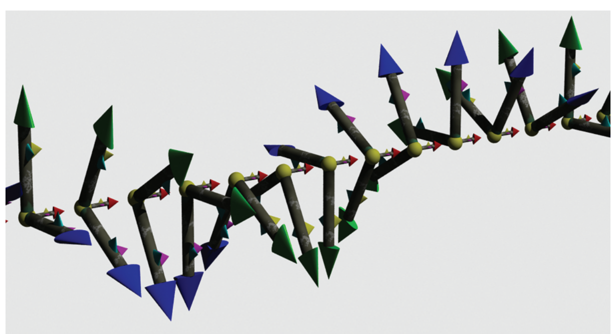

Figure 1. Representation of a single wavelength of a circularly polarised photon of equation (21). The electric field direction is represented using green arrowheads, the magnetic field blue and the momentum density red.

## 4. DISCUSSION

The new wave-function is consistent with the experimentally-observed field pattern for a photon. The extent that it truly describes a light quantum, a photon, is now discussed.

Firstly, consider that if equation (21) is a solution, then the linearity of field addition and the condition of energy conservation require that light comes in"lumps". Proof: consider an emission-absorption event in the frame of the photon. Now consider what happens if one superimposes a second such solution, where both overlap precisely in phase and wave-train length, such that twice the energy is transmitted at the same frequency in a single overlapping event. In this case the fields add everywhere. Since the energy density goes as the field squared, this would give four times the energy density everywhere and hence four times the energy transmitted. This violates energy conservation and hence such a process is, at least to first order,"not allowed". This is similar to the argument proposed in earlier work to explain the origin of the exclusion principle.[9] Viewing the same photon from a different frame scales energy density by $R^2$, but length by $R^{-1}$ leading to a linear increase in energy overall. The energy may be increased in a single event, self-consistent with both the linearity of energy and field and with relativity, is to increase the frequency. This is tantamount to varying the factor "$R$" in equation (21), affecting both the frequency and the overall energy by the same factor. This process gives a linear increase in both energy and field, as is required by the relativistic transformation of the solution and by experiment. This is the primary reason why the new first order, relativistic expression of equation (21) is necessarily physically quantised. Note that, conversely, relativity itself may be viewed as that transformation required to ensure the linearity of both energy and field as expressed by equation (19) and equation (20) above. Manifestly, for such wave-functions, the energy scales with frequency as this appears both in exponent and pre-factor. Clearly, this relation $h$ must be identified with the constant of Plank such that $\mathcal{E} = h\nu$. In other words, the Plank constant defines the scale of length for any given photon wave-function of the form of equation (21) with characteristic energy $\mathcal{E}$ . For a given proper energy, it sets the scale of rulers and clocks for that event. One may have different wave-lengths, but then one must also have correspondingly different energies - just as is observed. This is the key result of this paper: the quantisation of allowed solutions of the continuous theory is a consequence of the experimentally-observed conservation of energy and the linearity of field. It is worth noting that, to second order it may be possible to have double or triple photons. These would then have the quantisation rule $\mathcal{E} = hn^2\nu$, with $n$ a natural number. Such photons (Phat photons) may already have been observed, as has been discussed by Williams.[11]

Secondly, note that though the fact of quantisation of the kind of solution represented by equation (21) has not, here, required the introduction of a differential operator, a calculation of the value of the constant of proportionality between energy and frequency (Plank's constant) does. Charge appears at the level of the vector, and the field at its differential, bringing in a factor of R such as that in the pre-factor of equation (21). The question is then: can an expression be found relating the value of the elementary charge and that of Plank's constant in the present formalism? Such an estimate requires a study of the internal dynamics of the emitting and absorbing particles, at its simplest an electron, and this is beyond the scope of the present paper. In earlier work, however, a consideration of a simple semi-classical model of the electron as a localised photon did lead to such a relation.[10] This gives an estimate for Plank's constant in terms of the elementary charge in that model of $\hbar = 1.27 \times 10^{-34} Js$ which is, at least, of the right order of magnitude.

Thirdly, note that the field development and transformation parallel the Maxwell equations more closely than do more conventional solutions. The microscopic development of the field components is not merely a rotation. In equation (21), as one progresses forwards in space the field elements in the solution transform back and forth between electric and magnetic field components, just as in the case of equation (8) and equation (9). Although equation (22) looks just like a conventional simple electromagnetic wave-function, the underlying origin of the elements of electric and magnetic field, as described by eq. (21) is back and forth between each other, just as described by the Maxwell equations. Nonetheless, for a fixed position in some frame, as time progresses the field components appear to rotate (or oscillate), just as is the case for physical sources such as incandescent lights or a transmitter. In these respects, the new solutions match not only what is observed, but also parallel more closely the underlying field transformations of the Maxwell equations.

Fourthly, consider the relativistic transformations of space, time and field. By construction, the internal elements of space and time retain their proper relative form. This leads to the correct transformation properties of all components under a general Lorentz transformation.[6] In any other frame the proper relative transformation of the spatial and temporal field components is ensured in that they are constructed in such a way as to differ by the proper unit element from each other. In these solutions space and time transform properly with respect to each other. The result is that the transformed solution remains a solution in any proper Lorentz frame, right up to the limit of lightspeed. Put simply: all elements in the solution, both exponent (energy) and pre-factor (field) must scale linearly with R. Conversely, the Lorentz transformation is that transformation which ensures the linearity of both energy and field addition. Demanding that both energy and field should add linearly requires the introduction of the principle of absolute relativity. Athough the new principle was taken here as an ansatz, it appears that the principle of relativity of space and time is required in order to be consistent with the deeper principles of the conservation of energy and the linearity of the field. Absolute relativity is then not a postulate, but a requirement.

Fifthly, in any extension where $\omega$ and $k$ are not precisely equal, the separation between the space and time oscillations allows an identification with the two-phase harmony of de Broglie which lies at the root of quantum mechanics.[7,8] This corresponds to an extension to lightspeed quantum particles with rest mass, as has been discussed elsewhere.[9,10,12] The expansion to a full 4-dimensional wave-function introduces, necessarily, a limited extent perpendicular to the propagation direction. In an obvious extension, replacing the exponent with $(\alpha_1 kx + \alpha_2 ky + \alpha_3 kz - \alpha_0 \omega t)\alpha_{012}$ leads to the perpendicular $x$ and $y$ components having and expansion in terms of cosh and sinh instead of cos and sin. Explicitly, these are $F_0(\alpha_P \cosh(kx) - \alpha_{20} \sinh(kx))$ and $F_0(\alpha_P \cosh(ky) + \alpha_{10} \sinh(ky))$ respectively. These terms do not describe propagating solutions. Propagation is supported only along a line joining emitter and absorber and not transverse to the photon path. Further, both sinh and cosh functions increase exponentially in magnitude for larger lateral values, a clearly unphysical condition. $\mathcal{D}F = 0$ only if the expansion in the transverse direction is zero or constant. This confines the lateral, non-propagating dimension of the wave-function to lie close to the axis. In particular, some longitudinal components of field may be completely suppressed, since the sinh function is zero on axis. This may help to explain why the field of physical photons is primarily transverse. The scalar component, however, has a finite minimum on axis, and may supply a constant term. This term may prove to express the scalar mass-energy transferred by the photon from emitter to absorber.

Sixthly, if we demand further that the exponent should be consistent with a scalar wave-function, at least at the points of emission and absorption, such that it matches such wave-functions (of the form of equation (15), for example), then to achieve this, the factors in the exponent $(\alpha_3 kz - \alpha_0 \omega t)$ and $\alpha_{012}$ must commute. In particular, this requires that $\alpha_{012}$ commutes with the factor for the wavenumber $k$. This is only the case if that wavenumber corresponds to an integral number of half-wavelengths. This extra condition corresponds then to that of black-body quantisation.

Seventhly, consider simple transformations of the solution proposed in equation (21). Changing the sign of one component of the pre-factor alone, for example $(\alpha_{10} + \alpha_{31})$ to $(\alpha_{10} - \alpha_{31})$ has an interesting effect. This is no longer a left-handed solution for a wave propagating in the positive z direction, but a solution for a right-handed photon travelling in the negative z direction. That is, such a transformation matches precisely the physical process of reflection and the handedness of the field with respect to one another matches the direction of momentum transport. A change in the relative handedness of the electric and magnetic field components reverses the propagation direction (and flips the helicity). In other words, just as observed in experiment, the relative handedness of the electric and magnetic field components determines the direction of propagation. As discussed in the case of black-body quantisation above, changing the order of the unit angular momentum factor in the exponent from $(\alpha_3 kz - \alpha_0 \omega t)\alpha_{012}$ to $\alpha_{012}(\alpha_3 kz - \alpha_0 \omega t)$ is not a solution to the Maxwell equations. Indeed, for $k = \omega$ it is a frequency doubled oscillation. Though this is not a solution for the Maxwell equation, in the case of the electron-positron annihilation it does correspond precisely to the internal zitterbewegung frequency of the fermions as described by the Dirac equation.[3] It is tempting, then, to identify this double-frequency solution with an electron wave-function. This is not so, the solution remains too simple. For a description of a massive particle the solution must, at the very least, follow periodic boundary conditions such as those described in the simple semi-classical model considered in earlier work.[10] If this is done, this may give a description of purely electromagnetic charged particles with half-integral spin.[9] Simply changing the sign of the exponent remains a solution, but has the physical effect of transforming from left-handed to right-handed or vice-versa. Thus a linearly polarised photon may be represented as a sum or difference of such solutions. For the particular solution proposed, these are x polarised and y polarised respectively. Elliptical polarisations may be obtained from a linear combination in the usual way.

Finally, the wave-function in equation (21) describes a temporal rotation in real space. This means the lateral extent in the photon frame should not exceed a rotation horizon imposed by the speed of light. This imposes conditions on the angular momentum of allowed solutions. The concept was used in previous work to lay bare the physical origin of the anomalous magnetic moment of the electron as a localised photon.[10] For a given frequency the limit imposed by the speed of light on rotation about the photon axis, the rotation horizon, is just $r_h = c\omega$. Introducing the photon momentum observed in experiment, $\vec{p} = \hbar\omega/c$, gives a limit on the integral allowed angular momentum of the solutions of $r_h \times \vec{p} = \hbar$. This sets the intrinsic scale of unit angular momentum for solutions such as that described by equation (21). The form demanded by equation (17) and manifested in equation (21) is not merely descriptive, it is strongly proscriptive. Demanding the principle of absolute relativity, manifested in the form of equation (17), places strong restrictions on allowed solutions, over and above those required by the Maxwell equations alone. In summary: to be a travelling wave solution, the proper form must have electric fields perpendicular and of equal magnitude, must be associated with a unit angular momentum and energy must scale with frequency, just as is observed for the physical photon.

In conclusion, the properties discussed above mean that the new construction does not describe many non-physical combinations of propagating field, though it does allow those observed in nature. Equation (21) is only a solution if a strongly-constrained set of physical conditions are met. Crucial is that *both* the rate of change of phase *and* the field magnitude in both space and time scale with $R$. The magnitudes of the electric and magnetic field components must be equal, but are otherwise arbitrary. To be a solution of the free-space Maxwell equations, $\omega$ must equal $k$. The signs of $\omega$ and $k$ must match the handedness of the field components and the scale of frequency must match the scale of energy in the pre-factor. Further, a factor corresponding to a unit angular momentum is required in the exponent to transform the proper relative form of the spatial and temporal elements to a travelling wave solution with pure fields alone. In the logical extension of the 2D case to the 4D case the lateral components do not propagate. Once again, the required physical conditions match those observed in the physical photon.

## 5. EXPERIMENTAL TESTS

The theory is already consistent with a great number of classic experiments. The present theory, as discussed above, explains the underlying physical origin of effects such as the quantisation of light itself which must otherwise be taken simply from experiment. Equally, it explains why, though transmitted light appears quantised, local fields are simply continuous and well-described by classical electromagnetism. It has been argued that the present theory is more closely consistent with the whole existing body of experiment than are any alternatives. This does not, however, mean that it does does not suggest new avenues for experimental investigation.

The fact that the limits on the angular momentum of the photon may arise from the elementary charge opens up an interesting set of experimental possibilities. It is possible that photons emitted and absorbed by collective states of matter, such as in the superconducting or fractional quantum Hall regimes at low energy, or in regimes where fractional charges such as quarks may be present at high energy, may have different constraints on the limits of the photon angular momentum. It may therefore be possible to produce or detect photons with angular momentum a fraction or multiple of $\hbar$ with energy proportional to this fraction or multiple squared. This should be subject to experiment.

## 6. CONCLUSIONS

In conclusion, a stringent application of the principle of absolute relativity, demanding solutions of the form of equation (21), leads to solutions of the continuous first-order Maxwell equations with many of the properties of the physical photon. Such solutions propagate equal and perpendicular fields along an axis perpendicular to both. Perpendicular to the propagation axis they are strongly constrained. The transformation of the fields require that the energy transmitted should come in “lumps” and that this energy is proportional to the frequency. All such solutions have the same angular momentum, meaning that this appears quantised. An extra boundary condition at the point of emission and absorption requires a black-body quantisation. Taken together, it may be argued that these features mean that the new wave-function better represents the physical photon than do more conventional solutions.

### 6.1 Acknowledgments

This work flows from a long and fruitful collaboration with my friend and co-pirate Martin van der Mark. At least half of anything creditworthy in this article is due to him. I am grateful to Phil Butler for his education in algebra, to Stephen Leary for his work on the development of the wave-functions, to Tim Drysdale, John Duffield and David Mathes for useful suggestions and to Nick Green, Nick Bailey and John Weaver for their constant encouragement. The figure was drawn using Povray.

## REFERENCES

[1] Maxwell, J.C. A treatise on electricity and magnetism. 3rd edition, Dover(1954).

[2] Jackson, J.D. Classical Electrodynamics. (Wiley, New York, 1975).

[3] Dirac, P. A. M. The Principles of Quantum Mechanics, (Oxford U.P., London, 1958). 4th ed.

[4] Lounesto, P. *Clifford Algebras and Spinors*, *Lond. Math. Soc. Lecture Notes Series* **239**, (1997)

[5] Penrose, R. The Road to Reality, Jonathan Cape (2004).

[6] Leary, S.J. Investigation of electromagnetism in a real Dirac algebra. PhD Thesis, University of Glasgow, 2007.

[7] de Broglie, L. Waves and Quanta. *Nature* **112**, 540 (1923).

[8] de Broglie, L. Recherches sur la théorie des quanta, *Ann. Phys.* Ser. 10 **3**, 22 (1925).

[9] Williamson, J. G. 2012 Fermions from Bosons and the origin of the exclusion principle, Proceedings of MENDEL 2012

[10] Williamson, J.G. & van der Mark, M.B. Is the electron a photon with toroidal topology?, *Ann. Fondation L. de Broglie* **22**, 133 (1997).

[11] Williams, Pharis E. 2013 Phat photons and Phat lasers, Proc. SPIE 8832, The Nature of Light: What are Photons? V, 88320D (1 October 2013)

[12] Williamson, J. G. 2008 On the nature of the electron and other particles, paper presented at The Cybernetics Society 40th Anniversary Conference

[13] Williamson, J. G. A new theory of light and matter. Paper presented at FFP14, Marseille, France, July 2014.

[14] Williamson, J. G. The nature of the photon and the electron, companion paper to this one for August 2015.