The Dirac Equation for the Electron

April 13, 2007

Physics Theory Seminar Wesleyan University Middletown, Connecticut

Dr. Roy Lisker

OUTLINE

1. The Dirac equation for the electron. Translating its symbolism

- 2. Historical and physical motivations that led Paul Dirac to derive this equation in 1928
- 3. The Klein-Gordon equation
- 4. Why search for a linear version of the Klein-Gordon?
- 5. A digression into the theory of quadratic forms.
- 6. Dirac's factoring of the Klein-Gordon Equation
- 7. Consequences of the Dirac equation for physic

Note:

Before delivering this lecture I sent copies around to friends and colleagues for suggestions. The following is excerpted from the reply sent by a friend, a physicist living in the neighborhood of Bard College:

"Roy: Excellent exposition of the math.... Alas you have botched the history,,, Despite your efforts to the contrary, Dirac did visit Bard (twice)"..

(Comment: In the late 70's and early 80's I sent letters to socially designated distinguished intellectuals, including Isaac Bashevis Singer, Paul Dirac and Steven Smale, encouraging them to turn down invitations to speak from Bard College because it was a bad place. My letter may have had a contributing effect on Bashevis Singer's decision to turn down the College's invitation. Informants in the Physics Department let me know that Dirac had taken my letter seriously (We'd met briefly at the Einstein Centennial Symposium in 1979), and made his own investigation. Evidently he concluded (correctly) that academic fights of parochial local yokels were not a strong enough inducement to deprive the entire Hudson Valley region (that portion of it not banned from the Bard campus), of his presence and insight.

A similar letter to Leonard Bernstein helped lift his spirits after a particularly stupid article on him by Leon Botstein appeared in Harper's.

(The tragedy of Leonard Bernstein, May 1983) I know this because he called me up to thank me.)

"... we had some interesting conversations on the origins of Spinors and Pauli's role. His memory of 50-year old events was striking (as he couldn't remember lunch). He put his introduction to Cayley

elsewhere (in Engineering School at Bristol) and for other reasons (pragmatic problem solving –it turns out that Electrical Engineers, particularly English, actually read Heaviside rather than Gibbs, and linearizing (diagonalizing) expressions with circuit theory variables ('complex impedance' and such) was standard procedure.)

He slept through Pauli's lecture at Cambridge – he already had the spinors in hand.. He didn't worry about the classical electron radius –in fact he thought his anti-electrons were protons...

Not to worry. Mathematicians always make up fantasies about how physicists might have learned from them. Yours is at least intelligible, though Dirac is a bad target for this self-aggrandizement, as he was prone to gonzo mathematics. "

1. The Dirac equation for the electron. Translating its symbolism.

The Dirac Equation for the free (in the absence of an external field) electron is:

$$i \frac{\partial \psi}{\partial t} = (-i\alpha \cdot \nabla + m\beta)\psi$$

where

$$\boldsymbol{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{e\uparrow} \\ \boldsymbol{\psi}_{e\downarrow} \\ \boldsymbol{\psi}_{p\uparrow} \\ \boldsymbol{\psi}_{p\downarrow} \end{pmatrix}$$

has 4-components, called spinors, each of which satisfies the equation independently. The notation refers, respectively to "electron spin up" "electron spin down" "positron spin up" and "positron spin down".

 β is a 4-matrix given by

$$\beta = \begin{pmatrix} 0 & Id(2) \\ Id(2) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

 α is essentially a vector whose components are a version of the 3 Pauli spin matrices in a skewed 4th order matrix representation. That is to say:

$$\sigma = (\sigma_1, \sigma_2, \sigma_3)$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli spin matrices, and

$$\alpha = (\alpha_1, \alpha_2, \alpha_3)$$

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}; i = 1, 2, 3$$

These are actually 4-matrices in abbreviated form. Writing them out explicitly gives:

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}; \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}; \alpha_3 = \begin{pmatrix} 0 & 0 & 1 & o \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

These coefficients are elements of what is known as a Clifford Algebra. All of this will be explained in due course.

2..Historical and physical motivations that led Dirac to derive this equation in 1928

(a) Niels Bohr's model of the atom

The Bohr model for the atom is a familiar one. It is the picture of the atom as commonly accepted by the public, with a host of electrons circling around a central core or nucleus, consisting of a tight conglomerate of protons and neutrons.

This model was excellent for the description and prediction of atomic spectra, but quickly ran into problems, even in as simple a matter as the calculation of the number of electrons that can fill up places in a given orbital, or 'sub-orbit' specified by a given quantum number.

It was then shown by Wolfgang Pauli, that the Bohr model placed tight restrictions on the number of electrons that could occupy a given orbit. This is now known as the Pauli Exclusion Principle and applies to all fermions.

Although Bohr's theory required that there be only one electron per sub-orbit, or orbital, Pauli's calculations showed that there could be two. Pauli referred to a "classically non-describable duplexity", and it became customary to speak of a "duplexity paradox".

(b) The hypothesis of a spinning electron as an explanation for magnetism.

A solution to the duplexity paradox was proposed by the Dutch physicists Samuel Goudsmit and George Uhlenbeck in 1925. If the electron were also spinning as it traveled about its orbit, the

interaction with its electric field would transform it into a tiny bar magnet. This intrinsic magnetism would explain duplexity, because two and only two 'directions' of spin would be allowed, up and down.

(c)The conclusion that the electron would "tear itself apart" through rotation.

It was pointed out that, in order that the electron produce a magnetic field of the required strength, it would have to be spinning at an incredibly high speed. A calculation by Lorentz placed this speed at 137 times the speed of light! At such speeds the distribution of charge on the electron's surface would fly apart.

Finally there was a discrepancy in the geomagnetic ratio, g. This is the ratio of a charged particle's magnetism to its angular velocity of rotation. Classical calculations gave a value of g = 1. The theory of Goudsmit and Uhlenbeck postulated g = 2

Other contradictory properties of the original spin hypothesis proposed by Goudsmit and Uhlenbeck were discovered by Fermi and Rasetti

(d) Dirac's electrons as "point particles".

By 1928 several equivalent formulations of the principles of quantum theory had been developed, notably those of Heisenberg, based on matrices, Schrödinger based on his wave equation, and Dirac's transformation theory, based on extending classical Hamiltonian formalism to the quantum level by exploiting the analogy between Poisson Brackets {P,Q} and the Heisenberg commutator operator (P,Q) = P(Q) - Q(P). Dirac's transformation theory requires that one treat the electron as a point charge, with no volume or thickness. A point cannot be said to be spinning.

The impasse was resolved first, not by Dirac, but by Oscar Klein and, independently, Walter Gordon in 1926, through an elementary extension of the standard transformation scheme of the Schrödinger equation by Special Relativity. 3. The Klein-Gordon equation.

There is a formalism for transforming the classical Hamiltonian into the Schrödinger wave equation whereby one replaces observables such as energy, time, position and momentum, by operators.

Start with H = K + V

H is the Hamiltonian, numerically equal to the total energy K is kinetic energy, V potential. One way of writing the

kinetic energy is $K = \frac{1}{2m} \vec{p} \cdot \vec{p}$ where $\vec{p} = (mv_x, mv_y, mv_z)$ is the momentum. The Hamiltonian Equation becomes

$$H(\equiv E) = \frac{1}{2m}\vec{p} \bullet \vec{p} + V$$

Replace H by its numerical value, E, and treat V as a constant operator, that it is say, it multiplies whatever function is placed in front of it by its numerical value. In everything following we will make the convenient assumption that c=1, $h/2\pi = 1$

The replacement schema is:

$$E \to i \frac{\partial()}{\partial t}$$

$$\vec{p} \to -i\nabla()$$

$$V \to V \bullet ()$$

The "upside down delta" is a standard notation for the gradient.

When applying this to the Hamiltonian equation, one translates the dot product into composition of operators.

Formally,

which is the standard time-dependent Schrödinger wave equation.

The introduction of Special Relativity leads to important modifications of the observables of energy and momentum. Its formalism unites energy and momentum in a single 4-vector, in which energy appears at the time-component of the momentum. This can be notated in various ways. For example:

(1)
$$\vec{p}^{\mu} = (E, \vec{p}), \mu = 0, 1, 2, 3$$

(2) $p_{\mu}p^{\mu} = E^2 - \vec{p}^2$
(3) $E^2 = p^2 + m^2c^4$

(1) is the expression for the 4-momentum as a 4 dimensional vector.

(2) uses the Einstein convention of summing over the same repeated letter when it appears as both upper and lower index.

(3) is the relativistic energy, a more accurate version of $E = mc^2$. The letter c for the speed of light has been included for clarity.

What Klein and Gordon did was, simply to apply the transformation schema to equation (3):

$$E^{2} = p^{2} + m^{2}c^{4}$$

$$\downarrow$$

$$i \frac{\partial}{\partial t} (i \frac{\partial}{\partial t}) = -i\nabla(-i\nabla) + m^{2}$$

$$\therefore -\frac{\partial^{2}}{\partial t^{2}} \varphi = -\nabla^{2}\varphi + m^{2}\varphi$$

Once again one sets c =1. The symbol φ is used instead of ψ merely to indicate that this is a modification of the non-relativistic Schrödinger equation. This is the Klein-Gordon Equation.

As a footnote, let me add that there is a theorem, known as the *Groenwald-van Hove Theorem* which shows, by an explicit calculation, that the "Replacement Scheme" of Quantum Mechanics, whereby one replaces Observables and their products by Operators, breaks down when the order of the products exceeds 2. That is to say that even as expression as simple as the commutator of

(Position) ³ with (Momentum) ³ = $[Q^3, P^3]$, yields two different expressions when calculated in two equally valid ways. In other words, the measurement of Phase Space volume, which is the product of P³ with Q³, lies outside the formalism of Quantum Theory. For the details consult page 101, Sternberg and Guileman "Symplectic Techniques in Physics", Cambridge University Press, 1984.

4. Both Pauli and Dirac derived linearized reductions of the Klein-Gordon equation. Why was this deemed necessary?

(a) Dirac's transformation theory, based on the analogy of Poisson Brackets from Classical Mechanics with Heisenberg's quantum commutator, only works for linear equations.

(b) Born's interpretation of the square of the modulus of ψ as a probability, only works for linear equations.

(c) The Klein-Gordon equation has negative energy solutions for which it gives no explanation. This would not have mattered in the normal classical situation, in which one discards the solutions that don't fit. In the quantum situation however, the transitions from positive to negative energy states are an inevitable consequence of the probability interpretation of the wave function. One of the many ways in which Physics is distinguished from Mathematics is that mathematicians since Pythagoras have been enamored of quadratic expressions, whereas physicists prefer linear equations whenever possible. It is rare indeed to encounter a mathematical theory that tackles equations of the third or higher degrees, not to be confused with the number of dimensions, which can be anything. (*Just yesterday I came across a theorem about the close packing of 24-dimensional space by 24-dimensional spheres! However theequation of an n-sphere is a quadratic form in n variables.*)

A notable exception to this general rule is the recent proof in 1995 of Fermat's Theorem by Andrew Wiles, which uses the properties of the cubic polynomial $y^2 = ax^3 + bx + c$

A homogeneous quadratic form is a polynomial expression in n variables in which each component is of degree 2. For example:

$$Q(x.y.z) = 5x^2 + \pi xz - \sqrt{7}xy + y^2 + 8,529.1z^2$$

It is easily shown that any quadratic form can be converted into an expression consisting only of squares of the independent variables by a linear transformation, that is to say, a matrix A. If the coefficients are real, this matrix substitution can be so chosen that the coefficients of the new form Q'(x', y', z') will be 1, -1 or 0. Thus, the above form can be reduced to

$$Q(x.y.z) = \pm x^2 \pm y^2 \pm z^2$$

where I've not bothered to take the trouble to find out which combination of plus and minus signs will result.

An important theorem by Sylvester, the "inertia theorem", states that, in whatever fashion this reduction is made, the difference between the number of plus signs and the number of minus signs is an invariant. This is an essential feature of General Relativity. The "signature" of all metrics in the Riemannian spaces of General Relativity is +---, that is to say inertia = -2.

In physics, which deals with quantities and magnitudes, one often wants to factor higher order polynomials into linear factors. Lets examine this procedure systematically, starting from 1-variable expressions.

(1) $Q(x) = x^2$.

If x is a real variable, obviously one can factor Q into Q = q_1q_2 , where $q_1 = cx$, $q_2 = (1/c) x$, c being an arbitrary constant.

(2) $Q(x, y) = x^2 - y^2$.

This also can be factored in an obvious fashion as

$$Q(x, y) = (x - y)(x + y) = q_1 q_2$$

$$q_1 = c(x - y); q_2 = \frac{1}{c}(x + y)$$

The composition of linear operators is analogous in many respects to ordinary multiplication. Thus, the equation

$$\frac{\partial^2 f(x,t)}{\partial x^2} - \frac{\partial^2 f(x,t)}{\partial t^2} = 0$$

can be factored into two particular equations

$$\frac{\partial f(x,t)}{\partial x} - \frac{\partial f(x,t)}{\partial t} = 0$$
$$\frac{\partial f(x,t)}{\partial x} + \frac{\partial f(x,t)}{\partial t} = 0$$

to produce a general solution of the form:

$$z = A\varphi_1(x+t) + B\varphi_2(x-t)$$

(3) $Q(x, y) = x^2 + y^2$

It isn't possible to factor this over the field of the real numbers. In order to factor such expressions one must extend the field of real numbers **R**, to the field of complex numbers **C**. This was first done in the 16th century by the genius, doctor, astrologer and charlatan, Girolamo Cardano. However a real understanding of how to work with "imaginary" quantities, or complex numbers, did not emerge until the 18th century.

The 19th century saw the introduction of the idea of factoring over a field, or more generally, over some algebraic space, which means that the factors remain in the same space as the variables and functions in the original expression. Thus, the expression in (3) can be factored as

$$Q(x, y) = x^{2} + y^{2} = (x + iy)(x - iy)$$

= $z \bullet z^{*}$
 $z = c(x + iy); z^{*} = c(x - iy)$

Observe that the numbers x and y, as well as the quadratic form Q are all in the field of the real numbers R , but that the "new" numbers z, and z^* are in the "extended field" of the complex numbers. The constant c may be either real or imaginary.

It also turns out, and this is not trivial, that if x and y are replaced by complex numbers u and v from C, that Q can still be factored in the same way:

$$Q(u,v) = u^{2} + v^{2} = (u + iv)(u - iv) = (v + iu)(v - iu)$$

$$u = a + ib; v = c + id$$

$$u^{2} = a^{2} - b^{2} + 2iab; v^{2} = c^{2} - d^{2} + 2icd$$

$$u + iv = (a^{2} - b^{2} - 2cd) + i(2ab + c^{2} - d^{2})$$

Thus, the field "invented" for the factorization of quadratic forms in two real variables is "large enough" to permit factorization of quadratic forms put together from elements in that extended field. Complex numbers may have been considered mysterious when they were first discovered, but they lose much of their air of mystery when interpreted as matrices, invented by the English mathematician Cayley. We can represent i by a matrix in many ways. The simplest way is to write

$$i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; i^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2,$$

where I_2 is the identity for the semigroup of 2x2 real matrices. The solutions z and z^* of the quadratic form Q can then be written as

$$z = x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$
$$z^* = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$
$$zz^* = \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{pmatrix} x^2 + y^2 & 0 \\ 0 & x^2 + y^2 \end{pmatrix}$$
$$= (x^2 + y^2)I_2$$

The reason for demonstrating this relationship is to show how the study of ways to factor quadratic forms eventually became the search for algebras of matrices, in which these forms could be interpreted and factored. The next level shows that this step is inevitable.

(4)
$$Q(x, y, z) = x^2 + y^2 + z^2$$

It was quickly discovered that it is not possible to factor this expression over the field of the complex numbers. In fact, it can only be factored by going to the next level, that is to say quadratic forms in 4 variables, and setting one of the variables to 0. In order to understand how this is done, I have to say a few words, briefly, about the concept of a field. A field is the algebraic generalization of a space in which it's possible to do ordinary arithmetic. Keeping this in mind, its axioms are readily described:

A field F is an algebraic space that is closed under the operations of addition, multiplication, subtraction and division. Addition (+) is the operation of an Abelian group, that is the say a+b = b+a for elements a b in F. Multiplication (x) is also a group operation, with the qualification that the identity of addition, that is to say "0", does not have an inverse. Finally, addition and multiplication are related by the Distributive Law: If a, b and c are any 3 elements of F, then

a(b+c) = ab + ac

The real numbers are a field; the complex numbers are a field. The fractions p/q where p and q are integers and $q\neq 0$, are a field, Given any non-zero real number ω , one can construct a field by taking all polynomial expressions in ω , together with all ratios, sums and productions of these polynomials to make a field which is written as F(ω).

Not all fields can be represented by matrices. In fact the simplest fields, those of the integers modulo p, Zp, where p is a prime number, can't be represented by matrices. There also exist algebraic structures which arise naturally, which are not fields, and which can't be represented by matrices. Octonions are what is called a "division algebra". In it one can factor quadratic forms of up to 8 variables. Because its multiplication is non-associative one can't do much else.

(5)
$$Q(x, y, u, v) = x^2 + y^2 + u^2 + v^2$$

The search for a field over which this quadratic form can be factored was undertaken by William Rowan Hamilton, the same person after whom the "Hamiltonian" is named. He discovered a field H, the *quaternions*, over which it can be factored, provided that the variables are all real and the inertia is 4 (that is to say all components are positive). It was considered quite an innovation that the multiplication in this field is non-commutative. Nowadays it is generally understood that matrix multiplication need not be commutative.

It has been proven that the only continuous fields that can be represented by families of matrices with real numbers as entries are R, C, and H, that is to say, the reals, the complex numbers and the quaternions.

A quaternion is a linear expression involving one real variable and 3 "square roots" of minus 1 , i, j, and k. If q is an element of H it may be written as:

$$q = x + iy + ju + kv$$

The product rules for i, j and k are:

$$i^{2} = j^{2} = k^{2} = -1$$

 $ij = k = -ji$
 $jk = i = -kj$
 $ki = j = -ik$

Once again, these rules are easily understood when quaternions are represented as matrices.

The sums, products, ratios, etc. of 1, i, j, and k, with coefficients in the real numbers, generate the field H. The expression (5) may be factored over H as:

$$Q(x, y, u, v) = x^{2} + y^{2} + u^{2} + v^{2}$$

$$= (x + iy + ju + kv)(x - iy - ju - kv)$$

$$= qq^{*}$$

$$x + iy + ju + kv$$

$$x - iy - ju - kv$$

$$\overline{x^{2} + ixy + jxu + kxv}$$

$$-ixy + y^{2} - jiyu - kiyv$$

$$-jxu - ijyu + u^{2} - kjuv$$

$$-kxv - ikyv - jkuv + v^{2}$$

$$= x^{2} + y^{2} + u^{2} + v^{2}!$$

In this calculation it is assumed that all of the variables are real. It will be important when we come to the Dirac Equation to observe that this method doesn't work for the quadratic form:

(6)
$$Q(x, y, u, v) = x^2 + y^2 + u^2 - v^2$$
 !

Why is that so? One is tempted to rewrite $-v^2 as +(iv)^2$, then use the factorization over the quaternion field described above. The problem with this is that the "i" in expression (6) *is not the same "i" as the one that appears in the table of quaternions*. Indeed, this is a misnomer, and the quaternion "i" really ought to be replaced by another letter such as h . I've kept the standard notation only because of the difficulties involved in resisting tradition.

Exercise: Try substituting x, y, u and iv into the factorization qq* and see what happens.

6. Dirac's factoring of the Klein-Gordon Equation

The Klein-Gordon Equation is a second order linear operator equation, of type (6). It cannot be factored over the field of the quaternions. However it can be factored over a weaker algebraic space known as a Clifford algebra, invented by the English mathematician William Clifford and discovered independently by Paul Dirac in the 1920's. It is somewhat remarkable that all of the techniques involved in this subject were discovered or developed by English mathematicians or mathematical physicists: Cayley, Sylvester, Clifford, Hamilton and Dirac. The legend that Dirac rediscovered Clifford algebras by staring into a fireplace over a good dinner at an English pub may well be true, but it's clear that the context had prepared the way for him.

Here is a quote from "The Rainbow of Mathematics" by the historian of mathematics, Ivor Grattan-Guinness, 1997:

"Especially from the late 1880's, Heaviside revived the English liking for operator methods."

Let's re-examine the Klein-Gordon Equation:

$$\frac{\partial^2 \varphi}{\partial t^2} = \nabla^2 \varphi - m^2 \varphi$$
$$= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \varphi - m^2 \varphi$$

Expressed in a pure operator form, without the Schrödinger wave equation, and replacing the missing "c" for clarity, this becomes:

$$\frac{\partial^2}{\partial t^2} = \nabla^2 - m^2$$
$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - c^4 m^2$$

which is clearly of type (6). Since Dirac wanted a linear equation, he decided to start with a linear form of unknown coefficients, compose it with itself once, identify coefficients on both sides of the equation, and see if what he came up with made any sense! Thank goodness Dirac was not trained as a mathematician. By stumbling around in areas in which mathematicians knew how to go about doing things, he walked into paradises they might never have visited!

Thus, Dirac begins with an abstract form to which he will later try to ascribe some meaning:

$$i \frac{\partial \psi}{\partial t} = (-i\alpha \cdot \nabla + m\beta)\psi$$

The symbols α and β for the moment are meaningless letters, although α is going to be a "vector" of some kind = (α_1 , α_2 , α_3), and β will be a "scalar", so-called. Iterating both sides, one gets

$$i\frac{\partial}{\partial t}(i\frac{\partial}{\partial t}) = -\frac{\partial^2}{\partial t^2}$$
$$(-i\alpha \cdot \nabla + m\beta)(-i\alpha \cdot \nabla + m\beta)$$
$$=$$

The calculation on the right side is somewhat involved, but can be reduced to:

$$= -\sum_{1}^{3} \alpha_{i}^{2} \frac{\partial^{2}}{\partial x_{i}^{2}} - \sum_{1}^{3} (\alpha_{i}\alpha_{j} + \alpha_{j}\alpha_{i}) \frac{\partial^{2}}{\partial x_{i}\partial x_{j}} - im\sum_{1}^{3} (\alpha_{i}\beta + \beta\alpha_{i}) \frac{\partial}{\partial x_{i}} + \beta^{2}m^{2}$$

Comparing this with the Klein-Gordon equation one sees that the following relations have to hold between the alpha- and betacoefficients:

$$\alpha_i^2 = 1$$

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 0$$

$$\alpha_i \beta + \beta \alpha_i = 0$$

$$\beta^2 = 1$$

These are the defining relations for the matrices of Dirac's Clifford Algebra. By comparing them to the quaternions we will see that they bear a close resemblance but aren't the same. The relationship between them can be seen by looking at the way they relate to the Pauli matrices, which arise from Pauli's very different, yet equally ingenious linear reduction of the Klein-Gordon Equation. Pauli's technical feat leads to the two-component spinors for describing particles of spin 0, while Dirac's factorization leads to the 4-components spinors used in the description of spin 1/2 particles. Write the Pauli matrices as:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Since i is itself a matrix, one can expand these into 4th order matrices as

$$S_{1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$
$$S_{2} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$
$$S_{3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

The quaternions are related to the Pauli matrices by a simple product

$$q_{k} = i\sigma_{k}; k = 1, 2, 3$$

$$q_{1} = "i" = i\begin{pmatrix} 0 & 1\\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i\\ i & 0 \end{pmatrix}$$

$$q_{2} = "j" = i\begin{pmatrix} 0 & -i\\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1\\ -1 & 0 \end{pmatrix}$$

$$q_{3} = "k" = i\begin{pmatrix} 1 & 0\\ 0 & -1 \end{pmatrix} = \begin{pmatrix} i & 0\\ 0 & -i \end{pmatrix}$$

One sees how the use of the same letter "i" causes confusion. The matrix q₃ is quite different from the matrix for the square root of minus 1. Indeed the latter is included in the former!

In the 4th order representation, these become

$$Q_{1} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$
$$Q_{2} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$
$$Q_{3} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

We now compare these with the Clifford Algebra matrices that appear in the Dirac Equation.

$$\alpha_{1} = \begin{pmatrix} 0 & \sigma_{1} \\ \sigma_{1} & 0 \end{pmatrix}$$
$$\alpha_{2} = \begin{pmatrix} 0 & \sigma_{2} \\ \sigma_{2} & 0 \end{pmatrix}$$
$$\alpha_{3} = \begin{pmatrix} 0 & \sigma_{3} \\ \sigma_{3} & 0 \end{pmatrix}$$
$$\beta = \begin{pmatrix} 0 & I_{2} \\ I_{2} & 0 \end{pmatrix}$$

The σ 's of course are the Pauli matrices. As these expressions are abbreviations for 4 matrices. they can be also written as:

$$\begin{aligned} \alpha_1 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ \alpha_2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & -0 & 0 & 0 \end{pmatrix} \\ \alpha_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \\ \beta &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

The Clifford matrices do not generate a field. In fact, if one takes all productions and sums, with real number coefficients. of the Clifford matrices, one finds non-invertible, or singular matrices, and "divisors of 0", that is to say non-zero matrices which, when multiplied, give the 0-matrix as a product. This weaker structure is what is known as an "algebra", and it is not always possible to represent them with matrices. A recent example are the so-called "q-matrices" or "Quantum Deformed Matrices" which have become a growing field of mathematics, bringing together ideas from Quantum Field Theory and Knot Theory. 7. Consequences of the Dirac equation for physics. .

(1) Spinors and the Spinor Calculus. Pauli's method for factoring the Klein-Gordon equation led him to the discovery of 2component spinors. Dirac carried this further with the discovery of 4-component spinors. A mathematical theory was developed by Elie Cartan in which spinors can have any even number of components.

(2) *Anti-matter.* Dirac's original interpretation of the negative energy states posited a negative-energy field that filled all of space, with the exception of "holes" into which an electron could spontaneously vanish. Feynman proposed the existence of the positron which was discovered by Anderson in 1932

(3) *The probability current expression* derived from Dirac's equation describes the behavior of particles of spin ¹/₂.

(4) *Quantum Field Theory*, the many particle treatment of quantum theory which quantizes fields as well as observables, comes naturally out of the interpretation of Fermi- Dirac statistics, which are inherent in the Dirac equation.

