In: J.S.R. Chisholm/A.K. Commons (Eds.), **Clifford Algebras and their Applications in Mathematical Physics**. Reidel, Dordrecht/Boston (1986), 321–346.

CLIFFORD ALGEBRA AND THE INTERPRETATION OF QUANTUM MECHANICS

David Hestenes

ABSTRACT. The Dirac theory has a hidden geometric structure. This talk traces the conceptual steps taken to uncover that structure and points out significant implications for the interpretation of quantum mechanics. The unit imaginary in the Dirac equation is shown to represent the generator of rotations in a spacelike plane related to the spin. This implies a geometric interpretation for the generator of electromagnetic gauge transformations as well as for the entire electroweak gauge group of the Weinberg-Salam model. The geometric structure also helps to reveal closer connections to classical theory than hitherto suspected, including exact classical solutions of the Dirac equation.

1. INTRODUCTION

The interpretation of quantum mechanics has been vigorously and inconclusively debated since the inception of the theory. My purpose today is to call your attention to some crucial features of quantum mechanics which have been overlooked in the debate. I claim that the Pauli and Dirac algebras have a geometric interpretation which has been implicit in quantum mechanics all along. My aim will be to make that geometric interpretation explicit and show that it has nontrivial implications for the physical interpretation of quantum mechanics.

Before getting started, I would like to apologize for what may appear to be excessive self-reference in this talk. I have been pursuing the theme of this talk for 25 years, but the road has been a lonely one where I have not met anyone travelling very far in the same direction. So I will not be able to give much in the way of reference to the work of others, except on occasion when I found my road crossing theirs. I have reached a vantage point from which I can see where the road has been taking me pretty clearly. I will describe what I see so you can decide if you would like to join me on the trip.

Since the pursuit of my theme has been a personal Odyssey, I will supply a quasi-historical account of my travels to give you some sense of where the ideas came from and how they developed. I began a serious study of physics and mathematics only after a bachelors degree in philosophy and other meanderings in the humanities. Although that handicapped me in technical skills, which are best developed at an earlier age, it gave me a philosophical perspective which is unusual among American students of science. I had found in my studies of modern epistemology that the crucial arguments invariably hinged on some authoritative statement by the likes of Einstein, Bohr, Schroedinger and Heisenberg. So I concluded (along with Bertrand Russell) that these are the real philosophers, and I must scale the Olympus of physics to see what the world is really like for myself. I brought along from philosophy an acute sensitivity to the role of language in understanding, and this has been a decisive influence on the course of my studies and research.

Under the influence of Bertrand Russell, I initially believed that mathematics and theoretical physics should be grounded in Symbolic Logic. But, as I delved more deeply into physics, I soon saw that this is impractical, if not totally misguided. So I began to search more widely for a coherent view on the foundations of physics and mathematics. While I was a physics graduate student at UCLA, my father was chairman of the mathematics department there. This gave me easy access to the mathematics professors, students and courses. Consequently, I spent as much tlme on graduate studies in mathematics as in physics, I still regard myself as much mathematician as physicist.

I mention these personal details because I believe that influences from philosophy, physics and mathematics converged to produce a result that would have been impossible without any one of them. To be more specific about training that bears on my theme, from physics I became very skilful at tensor analysis because my mentor, Robert Finkelstein, was working on unified field theories, and I became familiar with the Pauli and Dirac algebras from courses in advanced quantum theory. In mathematics, I had one of the first courses in "modern" differential geometry from Barret O'Neill, and I studied exterior algebra and differential forms when the only good books on the subject were in French. Now this is what that background prepared me for:

Immediately after passing my graduate comprehensive examinations I was awarded a research assistantship with no strings attached, whereupon, I disappeared from the physics department for nearly a year. My father got me an isolated office on the fourth floor of the mathematics building where I concentrated intensively on my search for a coherent mathematical foundation for theoretical physics. One day, after about three months of this, I sauntered into a nearby math-engineering library and noticed on the "New Books shelf" a set of lecture notes entitled *Clifford Numbers and Spinors* by the mathematician Marcel Riesz [1]. After reading only a few pages, I was suddenly struck by the realization that the Dirac matrices could be regarded as vectors, and this gives the Dirac algebra a geometric meaning that has nothing to do with spin. The idea was strengthened as I eagerly devoured the rest of Reisz's lecture notes, but I saw that much would be required to implement it consistently throughout physics. That's what got me started.

About two months later, I discovered a geometrical meaning of the Pauli algebra which had been completely overlooked by physicists and mathematicians. I went excitedly to my father and gave him a lecture on what I had learned. The following is essentially what I told him, with a couple of minor additions which I have learned about since.

Physicists tacitly assign a geometric meaning to the Pauli matrices $\boldsymbol{\sigma}_k$ by putting them in one-to-one correspondence with orthogonal directions in Euclidean 3-space. The $\boldsymbol{\sigma}_k$ can be interpreted as unit vectors representing these directions, because their products have a geometric meaning. Thus, the orthogonality of $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ is expressed by the anticommutative product $\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 = -\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1$, which can also be regarded as Grassmann's outer product $\boldsymbol{\sigma}_1 \wedge \boldsymbol{\sigma}_2$, so the result can be interpreted geometrically as a directed area (a bivector). This implies a geometric meaning for the formula

$$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 = i\,, \tag{1.1}$$

which appears only as formal result in the textbooks on quantum mechanics. This formula tells us that i should be interpreted as the unit pseudoscalar for Euclidean 3-space, for it expresses i as a trivector formed from the outer product $\boldsymbol{\sigma}_1 \wedge \boldsymbol{\sigma}_2 \wedge \boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3$ of orthogonal vectors.

Equation (1.1) suggests that the unit imaginary $(-1)^{\frac{1}{2}}$ in quantum can be interpreted geometrically as the unit pseudoscalar i for physical 3-space, though, strictly speaking, i is related to $(-1)^{\frac{1}{2}}I$, where I is the identity matrix for the Pauli algebra. This idea turned out to be wrong, as we shall see. But the suggestion itself provided a major impetus to my research for several years. It demanded an analysis of the way the Pauli and Dirac algebras are used in physics.

Physicists generally regard the $\boldsymbol{\sigma}_k$ as three components of a single vector, instead of an orthonormal frame of three vectors as I have suggested they should. Consequently, they write

$$\boldsymbol{\sigma}\cdot\mathbf{a} = \sum_k \boldsymbol{\sigma}_k \cdot a_k \tag{1.2}$$

for the inner product of a vector $\boldsymbol{\sigma}$ with a vector $\mathbf{a}$ having ordinary scalar components a_k. To facilitate manipulations they employ the identity

$$\boldsymbol{\sigma}\cdot\mathbf{a}\,\boldsymbol{\sigma}\cdot\mathbf{b} = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})\,. \tag{1.3}$$

This is a good example of a redundancy in the language of physics which complicates manipulations and obscures meanings unnecessarily.

Formula (1.3) is a relation between two distinct mathematical languages, the vector algebra of Gibbs and the Pauli matrix algebra. This relation expresses the fact that the two languages have overlapping "geometric content," and it enables one to translate from one language to the other. However, by interpreting the $\boldsymbol{\sigma}_k$ as vectors generating a geometric algebra we can eliminate all redundancy incorporating both languages into a single coherent language. Instead of (1.2) we write

$$\mathbf{a} = \sum_k a_k \boldsymbol{\sigma}_k \tag{1.4}$$

expressing the expansion of a vector in terms of an orthonormal basis. Then (1.3) takes the form

$$\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + i(\mathbf{a} \times \mathbf{b}) \,, \tag{1.5}$$

where

$$\mathbf{a} \times \mathbf{b} = -i(\mathbf{a} \wedge \mathbf{b}) \tag{1.6}$$

defines the vector cross product as the dual of the outer product. By the way, this shows that the conventional distinction between polar and axial vectors in physics is really the distinction between vectors and bivectors in disguise, for (1.6) exhibits the vector $\mathbf{a}\times\mathbf{b}$ as a representation of the bivector $\mathbf{a} \wedge \mathbf{b}$ by its dual.

Sometime later I realized that the conventional interpretation of (1.2) as an inner product of $\mathbf{a}$ with a spin vector $\boldsymbol{\sigma}$ is completely consistent with my interpretation of the $\boldsymbol{\sigma}_k$ as a frame of vectors, for the physicists always sandwich the $\boldsymbol{\sigma}$ between a pair of spinors before calculations are completed. Thus, if ϕ is a 2-component Pauli spinor, one gets a spin density

$$\phi^\dagger \boldsymbol{\sigma}_k \phi = \tfrac{1}{2} \mathrm{Tr}\,(\phi\phi^\dagger \boldsymbol{\sigma}_k) = \rho \mathbf{s} \cdot \boldsymbol{\sigma}_k \,. \tag{1.7}$$

The notation on the right indicates that the matrix trace is equivalent to projecting out the components of a vector $\rho\mathbf{s}$ which is inherent in the matrix $\phi\phi^\dagger$ by dotting it with the basis vectors $\boldsymbol{\sigma}_k$. Thus, (1.2) gives us

$$\phi^\dagger \boldsymbol{\sigma}_k \cdot \mathbf{a}\, \phi = \sum_k \phi^\dagger \boldsymbol{\sigma}_k \phi a_k = \rho\, \mathbf{s} \cdot \mathbf{a} \,, \tag{1.8}$$

which is an ordinary inner product on the right.

These observations about the Pauli algebra reveal that is has a universal significance that physicists have overlooked. It is not just a "spinor algebra" as it is often called. It is a matrix representation for the geometric algebra $\mathcal{R}_3$, which, as was noted in my first lecture, is no more and no less than a system of directed numbers representing the geometrical properties of Euclidean 3-space. The fact that vectors in $\mathcal{R}^3$ can be represented as hermitian matrices in the Pauli algebra has nothing whatever to do with their geometric interpretation. It is a consequence of the fact that multiplication in $\mathcal{R}_3$ is associative and every associative algebra has a matrix represention. This suggests that we should henceforth regard the $\boldsymbol{\sigma}_k$ only as vectors in $\mathcal{R}_3$ and dispense with their matrix representations altogether, because they introduce extraneous artifacts like imaginary scalars.

I wondered aloud to my father how all this had escaped notice by Herman Weyl and John von Neumann, not to mention Pauli, Dirac and other great physicists who has scrutinized the Pauli algebra so carefully. When I finished my little talk, my father gave me a compliment which I remember word for word to this day, because he never gave such compliments lightly. He has always been generous with his encouragement and support, but I have never heard him extend genuine praise for any mathematics which did not measure up to his own high standards. He said to me, "you understand the difference between a mathematical concept and its representation by symbols. Many mathematicians never learn that."

My initial insights into the geometric meaning of the Pauli algebra left me with several difficult problems to solve. The first problem was to learn how to represent spinors in terms of geometric algebra without using matrices. Unfortunately, Marcel Riesz never published the chapter in his lectures which was supposed to be about spinors. From his other publications I learned that spinors can be regarded as elements of minimal left ideals in a Clifford algebra, and ideals are generated by primitive idempotents. But I had to find out for myself how to implement these ideas in quantum mechanics. So I spent much of the next three years intensively studying spinors and ideals in the Pauli algebra, the DIrac algebra and Clifford algebras in general. I was unaware that Professor Kähler was engaged in a similar study at about the same time. I could have profited from his publications [2,3], but I did not learn about them till more than a decade later.

The mathematical problem of constructing spinors and ideals for the Dirac and Pauli algebras is fairly simple. The real problem is to find a construction with a suitable geometrical and physical interpretation. Let me describe for you the solution which I developed in my doctoral dissertation [4]. As you know, the Dirac algebra is mathematically the algebra $C(4)$ of complex 4×4 matricies and this is isomorphic to the complex Clifford algebra $\mathcal{C}_4$. A physico-geometrical interpretation is imposed on $\mathcal{C}_4$ by choosing an orthonormal frame of vectors γ_μ (for $\mu = 0, 1, 2, 3$) to represent directions in spacetime. Of course, the γ_μ correspond exactly to the Dirac matrices in $C(4)$. In addition, I provided the unit imaginary in $\mathcal{C}_4$ with a physical interpretation by identifying it with the unit pseudoscalar of $\mathcal{R}_3$ as specified by (1.1). This entails a factorization of $\mathcal{C}_4$ into

$$\mathcal{C}_4 = \mathcal{R}_3 \otimes \mathcal{R}_2 \,. \tag{1.9}$$

Of course, this factorization is to be done so that the $\boldsymbol{\sigma}_k$ in $\mathcal{R}_3$ correspond to the same physical directions as the γ_k (for $k = 1, 2, 3$) in $\mathcal{C}_4$.

The factorization (1.9) implicitly identifies $\mathcal{C}_4$ with a real geometric algebra $\mathcal{R}_{p,q}$. As Lounesto [5] and others have noted, every complex Clifford algebra C_{2n} can be identified with a real algebra $\mathcal{R}_{p,q}$ where $p + q = 2n + 1$ and $n(2n + 1) + q$ are odd integers. These conditions on p and q imply that a unit pseudoscalar of $\mathcal{R}_{p,q}$ commutes with all elements of the algebra and has negative square. Therefore, it has the algebraic properties of the unit imaginary scalar in C_{2n}. In the case of physical interest $p + q = 4 + 1 = 5$, and we can choose an orthonormal basis $\mathrm{e}_0, \mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3, \mathrm{e}_4$ for $\mathcal{R}^{p,q}$ so the unit so the unit pseudoscalar can be written $i = \mathrm{e}_0\mathrm{e}_1\mathrm{e}_2\mathrm{e}_3\mathrm{e}_4$. From this we can see explicitly that

$$i^2 = \mathrm{e}_0^2\mathrm{e}_1^2\mathrm{e}_2^2\mathrm{e}_3^2\mathrm{e}_4^2 = -1 \,, \tag{1.10}$$

provided $q = 1, 3$ or 5. Thus, $\mathcal{C}_4$ is isomorphic to $\mathcal{R}_{4,1}$, $\mathcal{R}_{2,3}$ and $\mathcal{R}_{0,5}$. Among these alternatives the best choice is determined by geometrical considerations.

The simplest relation to the spacetime algebra $\mathcal{R}_{1,3}$ is determined by the projective conditions

$$\gamma_\mu = \mathrm{e}_\mu \mathrm{e}_4 \,. \tag{1.11}$$

Imposing the spacetime metric

$$\gamma_0^2 = -\mathrm{e}_\mu^2\mathrm{e}_4^2 = 1, \qquad \gamma_k^2 = -\mathrm{e}_k^2\mathrm{e}_4^2 = -1 \,,$$

we find that $\mathrm{e}_4^2 = \mathrm{e}_k^2 = -\mathrm{e}_0^2 = 1$. Thus, we should identify $\mathcal{C}_4$ with $\mathcal{R}_{4,1}$. You will recognize (1.11) as defining a projective map, whose general importance was pointed out in my first lecture. It identifies $\mathcal{R}_{1,3}$ with the even subalgebra of $\mathcal{R}_{4,1}$ as expressed by

$$\mathcal{R}_{4,1}^+ = \mathcal{R}_{1,3} \,. \tag{1.12}$$

To show that this is consistent with the factorization (1.9), we write

$$\boldsymbol{\sigma}_k = \gamma_k \mathrm{e}_0 = \mathrm{e}_k \mathrm{e}_{40} \,. \tag{1.13}$$

Then we find that the pseudoscalar for $\mathcal{R}_3$, $\mathcal{R}_{4,1}$ and $\mathcal{R}_{1,3}$ are related by

$$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 = i = \mathrm{e}_{01234} = \gamma_5\mathrm{e}_4\,, \tag{1.14}$$

where $\gamma_5 = \gamma_{01234} = \mathrm{e}_{01234}$. Thus, all the requirements for a geometrical identification of i have been met. It may be of interest also to note that (1.14) and (1.11) can be solved for

$$\mathrm{e}_4 = -i\gamma_5 \qquad \text{and} \qquad \mathrm{e}_\mu = i\gamma_5\gamma_\mu\,, \tag{1.15}$$

which expresses the generating vectors of the abstract space $\mathcal{R}_{4,1}$ in terms of quantities with direct physical interpretations.

I did not employ the projective mapping (1.11) in my dissertation, because I did not appreciate its significance until later, but what I did was equivalent to it for practical purposes. With a definite physico-geometrical interpretation for $\mathcal{C}_4$ in hand, I went on to the study of ideals and spinors in $\mathcal{C}_4$. As most of you know, the columns of a matrix are minimal left ideals in a matrix algebra, because columns are not mixed by matrix multiplication from the left. The Dirac matrix algebra $C(4)$ has four linearly independent minimal left ideals, because each matrix has four columns. The Dirac spinor for an electron or some other fermion can be represented in $C(4)$ as a matrix with nonvanishing elements only in one column, like so

$$\begin{bmatrix} \psi_1 & 0 & 0 & 0 \\ \psi_2 & 0 & 0 & 0 \\ \psi_3 & 0 & 0 & 0 \\ \psi_4 & 0 & 0 & 0 \end{bmatrix}, \tag{1.16}$$

where the ψ_i are complex scalars. The question arises: Is there a physical basis for distinguishing between different columns? The question looks more promising when we replace $C(4)$ by the isomorphic geometric algebra $\mathcal{R}_{4,1}$ in which every element has a clear geometric meaning. Then the question becomes: Is there a physical basis for distinguishing between different ideals?

The Dirac theory clearly shows that a single ideal (or column if you will) provides a suitable representation for a single fermion. This suggests that each ideal should represent a different kind of fermion, so the space of ideals is seen as a kind of fermion isospace. I developed this idea at length in my dissertation, classifying leptons and baryons in families of four and investigating possible interactions with symmetries suggested by features of the geometric algebra, including SU(2) gauge invariant electroweak interactions.

I was not very convinced by my own dissertation, however, because there was too much guesswork in associating ideals with elementary particles, though my theory seemed no less satisfactory than other theories around at the time. I was bothered even more by the relation (1.14) between the pseudoscalar i for $\mathcal{R}_3$ and γ_5 for $\mathcal{R}_{1,3}$, because the factor e_4 does not seem to make any sense in terms of spacetime geometry. I should add that this problem is inherent in conventional applications of the Dirac theory, as can be seen by rewriting (1.13) in the form

$$\gamma_k = \boldsymbol{\sigma}_k\mathrm{e}_0\,. \tag{1.17}$$

In conventional matrix representations this is expressed as the decomposition of a 4×4 matrix into a Kroenecker product of 2×2 matrices $\boldsymbol{\sigma}_k$ and e_0.

Before continuing with the story of my Odyssey, let me point out that I have been exploiting an elementary property of geometric algebras with general significance. Let $\mathcal{G}_n = \mathcal{R}_{p,q}$ be a geometric algebra generated by orthonormal vectors e_λ where $\lambda = 1, 2, \ldots, n$. Let $\mathcal{G}_{n-2}$ be the geometric algebra generated by

$$\sigma_k = \mathrm{e}_k\mathrm{e}_{n-1}\mathrm{e}_n \tag{1.18}$$

where $k = 1, 2, \ldots, n-2$, and let $\mathcal{G}_2$ be the geometric algebra generated by the vectors e_n and e_{n-1}. It is easy to see that the elements of $\mathcal{G}_{n-2}$ commute with those of $\mathcal{G}_2$. Therefore, $\mathcal{G}_n$ can be expressed as the Kroenecker product

$$\mathcal{G}_n = \mathcal{G}_{n-2} \otimes \mathcal{G}_2 \,. \tag{1.19}$$

This factorization has been known since Clifford, but it seems to me that its geometric meaning has been overlooked. So I want to emphasize that the vectors of $\mathcal{G}_{n-2}$ are actually trivectors in $\mathcal{G}_n$ with a common bivector factor. Geometrically, they determine a family of 3-spaces which intersect in a common plane.

It is important to realize that the decomposition (1.19) does not involve the introduction of any new kind of multiplication aside from the geometric product. I have employed the notation and terminology of the Kroenecker product only to emphasize that the factors commute. I would like to add that in publications since my thesis I have avoided use of the Kroenecker factorization (1.19) in order to emphasize geometric interpretation. However, I believe that the factorization will prove to be very important in Geometric Function Theory, because it reduces geometric functions to commuting factors which can be differentiated independently.

The Kroenecker decomposition (1.19) should be compared with the projective decomposition of a geometric algebra which I emphasized in my first lecture. These are two basic ways of relating geometric algebras with different dimensions, and I believe they should be employed systematically in the classification of geometric algebras and their properties. Most work on the classification of Clifford algebras ignore geometric considerations and develops a classification in terms of matrix representations. Without denying the value of such work, I suggest that a classification without matrices is desirable. This is in accord with the viewpoint of my first lecture that geometric algebra is more fundamental than matrix algebra.

2. REAL SPINOR REPRESENTATIONS

About six months after completing my doctorate, I found a way to resolve the problem of geometric interpretation which bothered me in my thesis. I had worked out detailed representations for Lorentz transformations and the equations of electrodynamics in terms of the real spacetime algebra $\mathcal{R}_{1,3}$ and separately in terms of the Pauli algebra before it dawned on me that the representations are identical if $\mathcal{R}_3$ and $\mathcal{R}_{1,3}$ are projectively related by

$$\boldsymbol{\sigma}_k = \gamma_k \gamma_0 \,, \tag{2.1}$$

so that

$$i = \boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2 \boldsymbol{\sigma}_3 = \gamma_0 \gamma_1 \gamma_2 \gamma_3 = \gamma_5 \,. \tag{2.2}$$

As we saw in my first lecture, this simplifies computations and makes perfect sense geometrically. It also eliminates the need to supplement spacetime with additional degrees of freedom in order to interpret the Dirac algebra geometrically, as in the extension to $\mathcal{R}_{4,1}$ which we just discussed. At first sight, however, this appears to be incompatible with quantum mechanics, because the imaginary scalar $i' = (-1)^{\frac{1}{2}}$ of the Dirac algebra appears explicitly in the Dirac equation, and, unlike the pseudoscalar $i' = \gamma_5$, it commutes with the γ_μ. To see how this difficulty can be resolved, we need to understand the representation of spinors in the real spacetime algebra $\mathcal{R}_{1,3}$.

The number of linearly independent minimal left ideals in $\mathcal{R}_{1,3}$ is two, half the number we found for the Dirac algebra. An orthogonal pair of such ideals is generated by the idempotents $\frac{1}{2}(1 \pm \boldsymbol{\sigma}_3)$, where $\boldsymbol{\sigma}_3 = \gamma_3 \gamma_0$ is a unit timelike bivector. Any multivector M in $\mathcal{R}_{1,3}$ can be written in the form

$$M = M_+ + M_- \,, \tag{2.3}$$

where

$$M_\pm = M\tfrac{1}{2}(1 \pm \boldsymbol{\sigma}_3) \tag{2.4}$$

are its component in the two orthogonal ideals. Since the entire algebra has $2^4 = 16$ components, each of the ideals has 8 components, exactly the number of real components in a Dirac spinor. Therefore, we should be able to write the Dirac equation for a spinor $\Psi = \Psi_+$ in one of the ideals. The "imaginary problem" is handled by observing that the ideals are invariant under multiplication not only on the left by any element of the algebra but also on the right by the pseudoscalar $i = \gamma_5$, since i commutes with $\boldsymbol{\sigma}_3$. To help clarify the correspondence with the Dirac theory, let me define right multiplication by i as a linear operator $\underline{i}$ by writing

$$\underline{i}\Psi = \Psi i\,. \tag{2.5}$$

This operator commutes with any operator multiplying Ψ from the left, for

$$\underline{i}(\gamma_\mu \Psi) = (\gamma_\mu \Psi) i = \gamma_\mu(\Psi i) = \gamma_\mu(\underline{i}\Psi)\,, \tag{2.6}$$

which justifies the operator equation $\underline{i}\gamma_\mu = \gamma_\mu \underline{i}$. Note that this is mapping of the associative law onto the commutative law.

Now the Dirac equation for a particle with mass m and charge e can be written in the form

$$\Box \Psi i - eA\Psi = m\Psi\,, \tag{2.7}$$

where $A = a_\mu \gamma^\mu$ is the electromagnetic vector potential. This looks just like the conventional Dirac equation when it is written in the form

$$\gamma^\mu(\underline{i}\partial_\mu - eA_\mu)\Psi = m\Psi\,, \tag{2.8}$$

but it employs the only real spacetime algebra $\mathcal{R}_{1,3}$ in which every element has a clear geometric-physical meaning. As soon as I understood that (2.7) really is equivalent to the Dirac equation, I salvaged what I could from my dissertation and quickly wrote a manuscript which was eventually published as the book *Space-Time Algebra* [6].

I learned much later that already in 1962 Kähler [3] had proposed a form of the Dirac equation with γ_5 operating on the right which is essentially equivalent to (2.7). He employed complex scalars, but the real and imaginary parts obey separate equations of the same form as long as the imaginary unit does not appear explicitly in the equation.

One thing bothered me about the Dirac equation in the form (2.7): What is the physical significance of the bivector $\boldsymbol{\sigma}_3$ which determines the ideal of the spinor? It took me nearly two years to answer that question. As described in Ref. [7], I found the answer by factoring into

$$\Psi = \psi U\,, \tag{2.9}$$

where ψ is an element of the even subalgebra $\mathcal{R}^+_{1,3}$, and

$$U = \tfrac{1}{2}(1 + \gamma_0)(1 + \boldsymbol{\sigma}_3)\,. \tag{2.10}$$

Note that

$$\gamma_0 U = U \tag{2.11}$$

and

$$i\boldsymbol{\sigma}_3 U = Ui\,. \tag{2.12}$$

Also, it can be proved that any ψ in $\mathcal{R}^+_{1,3}$ for which $\psi\widetilde{\psi} \neq 0$ can be written in the canonical form

$$\psi = (\rho e^{i\beta})^{\frac{1}{2}} R\,, \tag{2.13}$$

where ρ and β are scalars and $R\widetilde{R} = 1$. I should mention that the *reverse* $\widetilde{M}$ of any multivector M in $\mathcal{R}^+_{1,3}$ with k-vector parts M_k can be defined by

$$\widetilde{M} = M_0 + M_1 - M_2 - M_3 + M_4\,. \tag{2.14}$$

Note that (2.11) enables us to regard ψ as even in (2.9), because if ψ had an odd part, that part could be made even by multiplying it with γ_0 without altering (2.9).

The even multivector ψ is a new representation of a Dirac spinor, for which the Dirac equation takes the form

$$\Box\, \psi i\boldsymbol{\sigma}_3 - e\mathbf{A}\psi = m\psi\gamma_0\,, \tag{2.15}$$

as can be verified by multiplying it on the right by U to get back (2.7). At first this may look more complicated than the conventional form of the Dirac equation, but it gets much simpler as we learn to understand it, Ironically, I derived (2.15) in my book and promptly forgot about it for two years until I rederived it by a different method and discovered how ψ is related to the observables of the Dirac theory. In spite of the similarity of (2.7) and (2.15) to the Dirac equation, some people doubt that they are equivalent, so I should provide a proof. I will employ the complex Clifford algebra $\mathcal{C}_4$, which we know is isomorphic to the Dirac matrix algebra $C(4)$. A minimal ideal in $\mathcal{C}_4$ is generated by

$$U' = \tfrac{1}{2}(1+\gamma_0)(1 - i' i\boldsymbol{\sigma}_3) = \tfrac{1}{2}(1+\gamma_0)(1 - i'\gamma_5\gamma_3)\,, \tag{2.16}$$

where, as before, i'denotes the imaginary scalar and $i = \gamma_5$ is the pseudoscalar. Note that U' has the idempotent property $(U')^2 = 2U'$.

Also

$$\gamma_0 U' = U'\,, \tag{2.17}$$

and

$$i\boldsymbol{\sigma}_3 U' = \gamma_5\gamma_3 U' = i'U'\,. \tag{2.18}$$

Now it is obvious that every spinor Ψ' in the ideal generated by U' can be written in the form

$$\Psi' = \psi U'\,, \tag{2.19}$$

where ψ is an even multivector exactly as before. For any odd part ψ in (2.19) can be eliminated by using (2.17), and if i' appears explicity in any part of ψ, (2.18) allows us to replace it with $i\boldsymbol{\sigma}_3$. It is easy to find a matrix representation for U' which puts Ψ' in the column matrix form (1.16), so Ψ' is clearly equivalent to a Dirac spinor. Thus, we have established that Ψ, ψ and Ψ' are equivalent representations of a Dirac spinor. (See Appendix A of Ref. [8] for more details and a slightly different proof.) Now the proof that (2.15) is equivalent to the Dirac equation is trivial. We simply multiply it on the right by U' and use (2.17), (2.18) and (2.19) to get the conventional form of the Dirac equation

$$\gamma^\mu(i'\partial_\mu - eA_\mu)\Psi = m\Psi'\,. \tag{2.20}$$

The main import of this proof is that complex scalars do not play an essential role in the Dirac theory. As (2.18) indicates, the unit imaginary i' is replaced by a spacelike bivector $i\boldsymbol{\sigma}_3 = \gamma_2\gamma_1$, when we go from the complex to the real representation. This implies that a geometrical meaning for i' is implicit in the Dirac theory, for the bivector $i\boldsymbol{\sigma}_3$ is manifestly a geometrical quantity. This presents us with an important question to investigate: What is the physical significance of the *fact* that the unit imaginary in the Dirac theory represents a spacelike bivector?

To establish a physical interpretation for the spinor "wave functions" Ψ, ψ and Ψ', we need to relate them to "observables." I will simply assert an interpretation at the beginning, but by the time we are finished it will be clear that my interpretation agrees with the conventional interpretation of the Dirac theory. I begin with the interpretation of ψ, because it is simplest and most direct. Using $\widetilde{\psi} = (\rho e^{i\beta})^{\frac{1}{2}}\widetilde{R}$, the Dirac probability current is given by

$$J = \psi\gamma_0\widetilde{\psi} = \rho R\gamma_0\widetilde{R} = \rho v\,. \tag{2.21}$$

As shown in Ref. [6], R is the "spin representation" of a Lorentz transformation, so $v = R\gamma_0\widetilde{R}$ describes a Lorentz transformation of the timelike vector γ_0 into the "velocity vector" v. The factor ρ is therefore to be interpreted as a probability density. Note that the factor $e^{i\beta}$ disappears from (2.21) because i anticommutes with γ_0.

The spin (or polarization) vector of a Dirac particle is given by

$$\rho s = \psi\gamma_3\widetilde{\psi} = \rho R\gamma_3\widetilde{R}\,. \tag{2.22}$$

Strictly speaking the spin S is bivector quantity, but it is related to the vector s by

$$S = R(i\boldsymbol{\sigma}_3)\widetilde{R} = iR\gamma_3\widetilde{R}R\gamma_0\widetilde{R} = isv\,, \tag{2.23}$$

so one determines the other.

The vectors γ_0 and γ_3 in (2.21) and (2.22) are not necessarily related to the reference frame of any observer. They are singled out because they appear as constants in the Dirac equation. These constants are not seen in the conventional form of the Dirac theory because they are buried in an idempotent. Having exhumed them, we may attend to their physical meaning. The constant $i\boldsymbol{\sigma}_3 = i\gamma_3\gamma_0 = \gamma_2\gamma_1$ is especially important because of its role as the imaginary unit. Equation (2.23) shows that the wave function relates it to the spin.

Indeed, we could replace $i\boldsymbol{\sigma}_3$ by the spin S in the Dirac equation (2.15) by employing the identity

$$\psi i\boldsymbol{\sigma}_3 = S\psi\,. \tag{2.24}$$

This enables us to interpret a rotation generated by $i\boldsymbol{\sigma}_3$ as a rotation in the S-plane. The bivector $i\boldsymbol{\sigma}_3$ is also generator of electromagnetic gauge transformations. The Dirac equation (2.15) is invariant under a gauge transformation of the wave function replacing ψ by

$$\psi' = \psi e^{i\sigma_3\chi} = e^{S\chi}\psi \tag{2.25}$$

while A is replaced by $A' = A - \mathrm{e}^{-1}\Box\chi$. This transformation also leaves the spin and velocity vectors invariant, for $\psi'\gamma_0\widetilde{\psi}' = \psi\gamma_0\widetilde{\psi}$ and $\psi'\gamma_0\widetilde{\psi}' = \psi\gamma_3\widetilde{\psi}$. Indeed, we could define the electromagnetic gauge transformation as a rotation which leaves the spin and velocity vectors invariant.

Now we can give a detailed physical interpretation of the spinor wave function ψ in its canonical form (2.13). We can regard ψ as a function of 8 scalar paramenters. Six of the parameters are needed to determine the factor R which represents a Lorentz transformation. Five of these determine the velocity and spin directions in accordance with (2.21) and (2.22). The sixth determines the gauge or phase of the wave function (in the S-plane). We have already noted that ρ is to be interpreted as a probability density. The interpretation of the remaining parameter β presents problems which I do not want to discuss today, although I will make some observations about it later. As you know, this parameter is not even identified in the conventional formulation of the Dirac theory.

Having related ψ to observables, it is easy to do the same for Ψ and Ψ'. From (2.9) and (2.10), we find that $\Psi\widetilde{\Psi} = 0$, but

$$\Psi\gamma_0\widetilde{\Psi} = \psi(1+\gamma_0)\widetilde{\psi} = \rho\mathrm{e}^{i\beta} + J\,, \tag{2.26}$$

and

$$\Psi i\gamma_0\widetilde{\Psi} = \psi i\gamma_3(1+\gamma_0)\widetilde{\psi} = \rho\, i\, s(1+v\mathrm{e}^{i\beta})\,. \tag{2.27}$$

On the other hand, if we introduce the definition $\widetilde{i'} = -i'$ for the scalar imaginary, then (2.16) and (2.19) give us

$$\Psi'\gamma_0\widetilde{\Psi}' = \Psi'\widetilde{\Psi}' = \psi U^2\widetilde{\psi} = \psi 2U\widetilde{\psi} = \rho\mathrm{e}^{i\beta} + J - i'\rho\gamma_5 s(1+v\mathrm{e}^{i\beta})\,. \tag{2.28}$$

Equating corresponding k-vector parts of (2.26), (2.27) and (2.28) and calculating their components, we get the following set of equivalent expressions for the so-called "bilinear covariants" of the Dirac theory:

$$<\widetilde{\Psi}'\Psi'> = <\Psi\gamma_0\widetilde{\Psi}> = <\psi\widetilde{\psi}> = \ \rho\cos\beta, \tag{2.29a}$$

$$<\widetilde{\Psi}'\gamma_\mu\Psi'> = <\gamma_\mu\Psi\gamma_0\widetilde{\Psi}> = <\gamma_\mu\psi\gamma_0\widetilde{\psi}> = \ J\cdot\gamma_\mu, \tag{2.29b}$$

$$i'<\widetilde{\Psi}'\gamma_\mu\wedge\gamma_\nu\Psi'> = <\gamma_\mu\wedge\gamma_\nu\Psi\gamma_0\widetilde{\Psi}> = \ (\gamma_\mu\wedge\gamma_\nu)\cdot(\rho S\mathrm{e}^{i\beta}), \tag{2.29c}$$

$$i'<\widetilde{\Psi}'\gamma_5\gamma_\mu\Psi'> = <i\gamma_\mu\Psi\gamma_0\widetilde{\Psi}> = <\gamma_\mu\psi\gamma_3\widetilde{\psi}> = \ \rho s\cdot\gamma_\mu, \tag{2.29d}$$

$$<\widetilde{\Psi}'\gamma_5\Psi'> = <i\Psi\gamma_0\widetilde{\Psi}> = <i\psi\widetilde{\psi}> = \ -\rho\sin\beta. \tag{2.29e}$$

The angular brackets here mean scalar part, and I have used the theorem $<AB> = <BA>$ to put the terms on the left in the standard form of the Dirac theory. Equations (2.29b) and (2.29c) justify our earlier identifications of the Dirac current and the spin in (2.21) and (2.22).

We have completed the reformulation of the Dirac theory in terms of the real spacetime algebra $\mathcal{R}_{1,3}$. It should be evident that of the three different representations for a Dirac spinor, the representation ψ is the easiest to interpret geometrically and physically. So I will work with ψ exclusively from here on, with full confidence that its equation of motion (2.15) is 100% equivalent to the conventional Dirac equation. I suggest that we refer to ψ as the *operator representation of a Dirac spinor*, because it produces "ideal representations" by operating on idempotents and it produces observables by operating on vectors as in (2.21) and (2.22).

The most important thing we have learned from the reformulation is that the imaginary i' in the Dirac equation has a definite geometrical and physical meaning. It represents the generator of rotations in a spacelike plane associated with the spin. Indeed, we saw that i' can be identified with the spin bivector $\mathbf{S}$. I want to emphasize that this interpretation of i' is by no means a radical speculation; it is a fact! The interpretation has been implicit in the Dirac theory all the time. All we have done is make it explicit.

Clearly the identification of the imaginary i' with the spin bivector has far-reaching implications about the role of complex numbers in quantum mechanics. Note that it applies even to Schroedinger theory [7] when the Schroedinger equation is derived as an approximation to the Dirac equation. It implies that a degenerate representation of the spin direction by the unit imaginary has been implicit in the Schroedinger equation all along.

This is is the kind of idea that can ruin a young scientist's career. It appears to be too important to keep quiet about. But if you try to explain it to most physicists, they are likely to dismiss you as some kind of crackpot. The more theoretical physics they know the harder it is to explain, because they already have fixed ideas about the mathematical formalism, and you can't understand this idea without re-analyzing such basic concepts as how to multiply vectors. They quickly become impatient with any discussion of elementary concepts, so they employ the ultimate putdown: "What are the new predictions of your theory." If you can't come up with a mass spectrum or branching ratio, the conversation is finished. I learned early that you must be very careful when and where and how you voice such a crackpot idea.

Since I first publlshed the real spinor formulation of the Dirac theory in 1967, the only people (besides my students) who took it seriously enough to use it for something were the Frenchmen

Casanova [9], Boudet [10] and Quilichini [11]. I suppose that is partly because they are not conventional physicists. They found out some interesting things about the solutions of the Dirac equation for the hydrogen atom. But most physicists would be more concerned about how the formalism applies to quantum electrodynamics (QED). There they would encounter difficulties at once, which I suppose has induced some to dismiss the entire formalism. Conventional QED calculations involve multiplications with a commuting imaginary all over the place, but the real formalism does not contain such an entity. It appears at first that one cannot even define the conventional commuting momentum and spin projection operators without it. I must admit that I was perplexed by this difficulty for a long time myself. However, it has a simple solution: One merely handles right multiplications by defining them as operators acting from the left, as I did in defining $\underline{i}$ by (2.5). Let me illustrate this by defining projection operators for the real spinor Ψ. The momentum projection operators are just the usual idempotents

$$P_\pm = \frac{1}{2m}(1 \pm p)\,, \tag{2.30}$$

where $p^2 = m^2$. The spin projection operators are defined by

$$\Sigma_\pm = \tfrac{1}{2}(1 \pm \underline{i}\gamma_5 s)\,, \tag{2.31}$$

where s is a spacelike unit vector satisfying $p \cdot s = 0$. Operating on Ψ this gives

$$\Sigma_\pm \Psi = \tfrac{1}{2}(1 \pm \underline{i}\gamma_5 s)\Psi = \tfrac{1}{2}(\Psi \pm i\, s\Psi i)\,. \tag{2.32}$$

Note that $\Sigma_\pm$ is not an idempotent of the algebra $\mathcal{R}_{1,3}$; nevertheless, it has the idempotent property $(\Sigma_\pm)^2 = \Sigma_\pm$ of a projection operator. Also, the operators $P_\pm$ and $\Sigma_\pm$ commute, like they do in the conventional theory.

One can apply the same trick to carry out conventional QED calculations with the spinor Ψ. However, the use of Ψ suggests new methods of calculation without projection operators which may prove to be superior to conventional methods. But that is a long story itself which I cannot get into today.

Now that we have settled the relation of the real spinor formalism to the conventional complex spinor formalism (I hope), we are prepared to discuss implications for the interpretation of quantum mechanics. First let me point out some negative implications. The complex probability amplitudes that appear in physics have led some to propose that quantum mechanics entails some new kind of "quantum logic" which is essentially different from the logic of classical thoery. However, the association of the unit imaginary with the spin shows that complex amplitudes arise from some physical reason, so the quantum logic idea appears to be on the wrong track.

Another negative implication of our reformulation concerns the interpretation of operators in quantum mechanics. It is widely believed by physicists that the Pauli and Dirac matrices have some special quantum mechanical significance, so their commutation relations have bearing on questions about quantum mechanical measurement. But the reformulation in terms of spacetime algebra shows that these commutation relations express geometrical relations which are no more quantum mechanical than classical. Therefore, we can dismiss most of that stuff as arrant nonsense. It has validity only to the extent that it is merely an expression of geometrical relations. For example, the γ_μ are often said to be velocity operators in the Dirac theory. But we see them merely as ordinary vectors, which are velocity operators only in the sense that they pick out velocity components from the wave function by the ordinary inner product, as, in fact, they do in (2.29b). To attribute more meaning than that to the γ_μ is to generate nonsense. There is, indeed, an extensive literature on such nonsense. We should not be surprised that this literature is muddled and barren.

Turning now to the positive implications of the real reformulation, we have seen that it reveals geometrical features of the Dirac theory which are hidden in the conventional matrix formulation.

Of course, it cannot *per se* produce any new predictions, because the reformulated Dirac theory is isomorphic to the original formulation. We will get new physics only if the new geometrical insights guide us to significant modifications or extensions of the Dirac theory. I want to tell you next about some promising possibilities along this line.

3. CLASSICAL SOLUTIONS OF THE DIRAC EQUATION

The relation of the Dirac theory to classical relativistic electrodynamics is not well understood. I aim to show you that it is more intimate than generally suspected. The classical limit is ordinarily obtained as an "eikonal approximation" to the Dirac equation. To express that, in our language, the wave function is put in the form

$$\psi = \psi_0 e^{i\sigma_3 \chi}\,, \tag{3.1}$$

and the "amplitude" ψ_0 is assumed to be slowly varying compared to the "phase" χ, so it is a good approximation to neglect derivatives of ψ_0 in the Dirac equation. Thus, inserting (3.1) into the Dirac equation (2.15), multiplying by $\widetilde{\psi}$ on the right, and using the canonical form (2.13) for ψ, we obtain

$$\Box\chi + eA = -mve^{i\beta}\,. \tag{3.2}$$

This implies that $e^{i\beta} = \pm 1$, because the trivector part on the right must vanish. These two signs can be absorbed in the sign of the charge e on the left. That tells us that the parameter β distinguishes between particle and antiparticle solutions of the Dirac equation. Assuming $\beta = 0$, by squaring (3.2) we obtain

$$(\Box\,\chi + eA)^2 = m^2\,. \tag{3.3}$$

This is exactly the classical relativistic Hamilton-Jacobi equation for a charge particle. On the other hand, the curl of (3.2) gives us

$$-m\Box \wedge v = \Box \wedge A = eF\,, \tag{3.4}$$

where $F = \Box \wedge A$ is the electromagnetic field. Now we use the general identity $v \cdot (\Box \wedge v) = v \cdot \Box v - \frac{1}{2}\Box v^2$. Since $v^2 = 1$ here, this implies that $v \cdot (\Box \wedge v) = v \cdot \Box v = \dot{v}$, where the overdot indicates a derivative along the streamlines (integral curves) of the vector field $v = v(x)$. Let me remind you that the probability conservation law $\Box \cdot J = \Box \cdot (\rho v) = 0$ implies that these curves are well-defined. Thus, by dotting (3.4) with v, we obtain the equation of motion for any streamline of the Dirac current,

$$m\dot{v} = eF \cdot v\,. \tag{3.5}$$

You will recognize this as the classical equation of motion for a point charge.

All this is supposed to be an approximation to the Dirac theory. But I want to point out that it holds exactly for solutions of the Dirac eauation when

$$\Box\psi_0 = 0\,, \tag{3.6}$$

in which case our assumption that derivatives of ψ_0 are negligible is unnecessary. Members of this audience will recognize (3.6) as a generalization of the Cauchy-Riemann equations to spacetime, so we can expect it to have a rich variety of solutions. The problem is to pick out those solutions with physical significance. To see how that can be done, we write $\psi_0 = \rho^{\frac{1}{2}} R$ with $v = R\gamma_0\widetilde{R}$, from which we obtain $v\psi_0 = \psi_0\gamma_0$. Differentiating, we have

$$\Box\,(v\psi_0) = (\Box v)\psi_0 - v(\Box\psi_0) + 2v \cdot \Box\psi_0 = \Box\psi_0\gamma_0\,.$$

Then using (3.6) and $\Box \cdot (\rho v) = \rho\Box \cdot v + v \cdot \Box\rho = 0$, we obtain an equation of motion for the spinor R along a streamline.

$$\dot{R} = -\tfrac{1}{2}(\Box \wedge v)R\,. \tag{3.7}$$

When (3.4) is used to eliminate $\Box \wedge v$, this becomes

$$\dot{R} = \frac{e}{2m} FR\,, \tag{3.8}$$

which, as we saw in my first lecture, can regarded as a classical equation of motion. And from the first lecture we know that besides implying the classical Lorentz force (3.5), it determines the precession of the spin vector along a streamline.

Our derivation shows that equations (3.6) and (3.8) are related by the probability conservation law. Any solution of the Dirac equation which satisfies these equations can fairly be called a classical solution. Let me tell you that exact classical solutions of the Dirac equation actually exist, though I do not have time to spell out the details here. The so-called Volkov solution for an electron in an electromagnetic plane wave field is one. It may be that there are solutions of this type for any electromagnetic field. However, the standard solution for the hydrogen atom is not one of them, for the parameter β in that case is a nonvanishing, nontrivial function of position. Nevertheless, it must have a definite relation to the known classical solution of (3.8) for the hydrogen atom. More work is needed along this line.

The intimate relation between streamlines of the Dirac theory and trajectories of the classical theory which we have just uncovered provides a much more detailed correspondence between the classical and quantum theories than the conventional approach using expectation values and Ehrenfest's theorem. Let me remark in passing that it also suggests a relativistic generalization of the Feynmann path integral including spin, where one integrates the entire classical spinor R along each path instead of just the gauge factor $\exp{(i\boldsymbol{\sigma}_3\chi)}$, where χ is the classical action. But what I want to emphasize most is that the basic idea which we have been exploiting provldes a general geometrical approach to the interpretation of the Dirac theory as follows: Any solution $\psi = \psi(x)$ of the Dirac equation (2.15) with the form (2.13) determines a field of orthonomal frames $\mathrm{e}_\mu = \mathrm{e}_\mu(x)$ defined by

$$\psi\gamma_\mu\widetilde{\psi} = \rho\mathrm{e}_\mu, \qquad \text{where} \qquad \mathrm{e}_\mu = R\gamma_\mu\widetilde{R}\,, \tag{3.9}$$

with $\mathrm{e}_0 = v$ and $\mathrm{e}_3 = s$ as before. Through each spacetime point there is a streamline $x = x(\tau)$ with tangent $v = v(x(\tau))$, and we can regard $\mathrm{e}_\mu = \mathrm{e}_\mu(x(\tau))$ as a "comoving frame" on the streamline with vectors e_1 and e_2 rotating about the "spin axis" $\mathrm{e}_3 = s$. For the classical solutions discussed above, the general precession of the comoving frame is determined by (3.8), while an additional rotation of e_1 and e_2 is determined by the "gauge factor" in (3.1). It should be of genuine physical interest to identify and analyze any deviations from this classical rotation which quantum mechanics might imply.

The Dirac theory provides a beautiful mathematical theory of spinning frames on the spacetime manifold. But a spinning frame is not a spinning thing, and physicists want to know if the Dirac theory can be interpreted as a mathematical model for some physically spinning thing. I am afraid that question cannot be answered without becoming embroiled in speculations. But the question is too important to be avoided for that. Some physicists have attempted to model the electron as a small spinning ball. But that introduces all kinds of theoretical complications and, as far as I have been able to see, no significant insights. Along with Asim Barut and others, I think a much more promising possibility is the idea that the electron is a particle executing a minute helical motion, called the zitterbewegung, which is manifested in the electron spin and magnetic moment. As I have recently published a speculative article on that idea [12], I will not go into details here. I only want to mention a key idea of that article, namely, that the Coulomb field ordinarily attributed to an electron is actually the time average of a more basic periodic electromagnetic field oscillating with the de Broglie frequency $\omega = mc^2/\hbar \approx 10^{21}\mathrm{s}^{-1}$ of the electron. This is a new version of wave-particle duality, where the electron is a particle to which this high frequency electromagnetic field (or wave) is permanently attached. As the article points out, this gives us a mechanism for explaining the most perplexing features of quantum mechanics from diffraction to the Pauli principle. What I want to add here is that this version of wave-particle duality may be viable even without the literal electron zitterbewegung. In particular, it may fit into the gauge theory which I turn to next.

4. GAUCE STRUCTURE OF THE DIRAC THEORY

One thing that bothered me for a long time after I discovered the underlying geometrical structure of the Dirac theory is the question: How does the concept of probability fit in with all this geometry? I now have an answer which I find fairly satisfying. To get a spacetime invariant generalization of the concept of a continuous probability density in 3-space, we need to introduce the concept of a probability current J obeying the conservation law $\Box \cdot J = 0$. Now, it should be evident from our earlier discussion that *every timelike vector field* $J = J(x)$ can be written in the form

$$J = \psi\gamma_0\widetilde{\psi}\,, \tag{4.1}$$

where $\psi = \psi(x)$ is an even spinor field. This theorem automatically relates the probability current to a spinor field and all the geometry that goes with it. You will notice that the bilinear relation between the spinor field ψ and the vector field J is perfectly natural in this language, and it is as applicable to classical theory as to quantum mechanics. This removes much of the mystery from the bilinear relation of the wave function to observables in quantum mechanics. It shows that the bilinearity has a geometrical origin, which certainly has nothing to do with any sort of quantum logic.

The relation of vector J to spinor ψ in (4.1) cannot be unique, because only four scalar parameters are required to specify J uniquely, while eight are required to specify ψ. Indeed, the same J results from (4.1) if ψ is replaced by

$$\psi' = \psi S \tag{4.2}$$

where

$$S\gamma_0\widetilde{S} = \gamma_0\,. \tag{4.3}$$

This is a gauge transformation of the wave function ψ, so the set of all such transformations is the *gauge group of the Dirac probability current*, that is, the group of gauge transformations leaving the Dirac current invariant. We can identify the structure of this group by decomposing S into

$$S = e^{i\alpha}U\,, \tag{4.4}$$

where α is a scalar parameter and $U\widetilde{U} = 1$. The set of all $S = U$ satisfying (4.3) is the spin "little group" of Lorentz transformations leaving the timelike vector γ_0 invariant; it has the SU(2) group structure. It follows from (4.4), therefore, that the gauge group of the Dirac current has the SU(2) $\otimes$ U(1) group structure. Note that this is a 4-parameter group, so it accounts completely for the difference in the number of parameters needed to specify J and ψ. You will notice also that it has the same structure as electroweak gauge group in the Weinberg-Salam (W-S) model of weak and electromagnetic interactions. I want to make the stronger claim that the *electroweak gauge group should be identified with the gauge group of the Dirac current.* This, of course, is to claim that the electroweak gauge group has been inherent in the Dirac theory all the time, though one could not see it without the reformulation in terms of spacetime algebra. It is a strong claim, because it relates the electroweak gauge group to spacetime geometry of the wavefunction. That requires some justification.

The gauge group of the Dirac current has a subgroup which also leaves the spin vector $\rho s = \psi\gamma_3\widetilde{\psi}$ invariant, namely, the 2-parameter group of elements with the form

$$S = e^{i\alpha}e^{i\sigma_3\chi}\,. \tag{4.5}$$

We recognize the last factor as an electromagnetic gauge transformation, which, as we have seen, leaves the Dirac equation as well as the spin and velocity vectors invariant. Now, the main idea of the W-S model is to generalize the electromagnetic gauge group of the Dirac theory to account for weak interactions. Since we have proved that the generator of electromagnetic gauge transformations is a bivector $i\boldsymbol{\sigma}_3$ associated with the spin, we should not be satisfied with a generalization which does not supply an associated geometric interpretation for the generators of weak gauge transformations. But we have just what is needed in the gauge group of the Dirac current. This appears to be ample justification for identifying that group with the electroweak gauge group. Of course, that requires a modification of the Dirac equation to accommodate the larger gauge group. As I have shown elsewhere [13] how that can be done to conform perfectly to the W-S model, I will not go into such details here. But I want to add some observations about the general structure of the theory.

An important feature of the W-S model is the separation of "right-handed" and "left-handed" components of the wave functions. To show how the conventional expressions for these components should be translated into our geometrical language, it is convenient to introduce the imaginary operator i defined by

$$\underline{i}\psi = \psi i\boldsymbol{\sigma}_3 \,. \tag{4.6}$$

Now the projections onto left- and right-handed components ψ_+ and ψ_- are defined by

$$\tfrac{1}{2}(1 \pm \underline{i}\gamma_5)\psi = \tfrac{1}{2}(\psi \pm i\psi i\boldsymbol{\sigma}_3) = \psi\tfrac{1}{2}(1 \mp \boldsymbol{\sigma}_3) = \psi_{\mp} \,. \tag{4.7}$$

Thus, we are back to spinors in ideals again, but now we have a physical interpretation for the ideals.

In my formulation of the W-S model in terms of spacetime algebra [13], the two ideals are interpreted as electron and neutrino eigenstates of the lepton wave function. Equation (4.7) shows that there is a relation between the representation of different particles by the two ideals and the decomposition of the wave function for one particle into left- and right-handed components. This is a relation between two features introduced as independent assumptions in the original W-S model. Moreover, from (4.5) we see that the quantities

$$i\mathbf{Q}_{\pm} = \tfrac{1}{2}(1 \pm \boldsymbol{\sigma}_3)i \tag{4.8}$$

are generators of the "spin invariance group" (4.5). This relates the electromagnetic gauge group to the "chiral projection operators" in (4.7). All this suggests possibilities for a deeper geometrical justification of the phenomenological W-S model. I should mention that the full W-S model involves a spinor with both even and odd parts, whereas (4.7) is an even spinor only. That must be taken into account in relating (4.7) to the W-S model.

For the purpose of physical interpretation we need to relate the decomposition (4.7) to observables. Since $(1 + \boldsymbol{\sigma}_3)\gamma_0(1 - \boldsymbol{\sigma}_3) = 0$, the ideals are orthogonal in the sense that

$$\psi_+\gamma_0\widetilde{\psi}_- = \psi_-\gamma_0\widetilde{\psi}_+ = 0 \,. \tag{4.9}$$

This implies that the Dirac current separates into uncoupled left and right handed currents $J_{\pm}$. Thus, using $\psi = \psi_+ + \psi_-$, we find

$$J = \psi\gamma_0\widetilde{\psi} = J_+ + J_- \,., \tag{4.10}$$

where

$$J_{\pm} = \psi_{\pm}\gamma_0\widetilde{\psi}_{\pm} = \psi\tfrac{1}{2}(\gamma_0 \pm \gamma_3)\widetilde{\psi} \,. \tag{4.11}$$

The expression on the right of (4.11) follows from

$$\tfrac{1}{2}(1 \pm \boldsymbol{\sigma}_3)\gamma_0\tfrac{1}{2}(1 \mp \boldsymbol{\sigma}_3) = \tfrac{1}{2}(1 \pm \gamma_3\gamma_0)\gamma_0 = \tfrac{1}{2}(\gamma_0 \pm \gamma_3) \,.$$

The corresponding decomposition of the spin vector is given by

$$\rho s = \psi\gamma_3\widetilde{\psi} = \rho(s_+ + s_-)\,, \tag{4.12}$$

where

$$\rho s_\pm = \psi_\pm\gamma_3\widetilde{\psi}_\pm = \psi\tfrac{1}{2}(\gamma_3 \pm \gamma_0)\widetilde{\psi} = J_\pm\,. \tag{4.13}$$

The separation (4.10) into uncoupled currents suggest that the two spin components $\psi_\pm$ might be identified with different particles, or coupled differently to gauge fields, as, in fact, they are in the W-S theory. Equation (4.11) shows that the currents $J_\pm$ are null vectors, as required for massless particles. And (4.12) shows that spin vectors $\rho s_\pm$ are, respectively, parallel or antiparallel to their associated currents $J_\pm$. I believe these elementary observations will be important for understanding and assessing the geometric structure of the W-S model.

5. CONCLUSION

My objective in this talk has been to explicate the geometric structure of the Dirac theory and its physical significance. My approach may seem radical at first sight, but I hope you have come to recognize it as ultimately conservative. It is conservative in the sense that, by restricting my mathematical language to the spacetime algebra, I allow nothing in my formulations of physical theory without an interpretation in terms of spacetime geometry. I am not opposed to investigating the possibilities for unifying physical theory by extending spacetime geometry to higher dimensions, and I believe geometric algebra is the ideal tool for that. But we still have a lot to learn about the physical implications of conventional spacetime structure, so I have focused my attention on that.

REFERENCES

[1] Marcel Riesz,'Clifford Numbers and Spinors', *Lecture series* No. **38**, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland (1958).

[2] E. Kähler, *Abh. Dt. Akad. Wis. MPTJ* **5** (1961).

[3] E. Kähler, *Rend. Mate. (Roma) Ser.* V, **21**, 425 (1962).

[4] D. Hestenes, *Geometric Calculus and Elementary Particles* (doctoral dissertation) UCLA (1963).

[5] P. Lounesto, *Found. Phys.* **11**, 721 (1981).

[6] D. Hestenes, *Space-Time Algebra*, Gordon and Breach, N.Y. (1966).

[7] D. Hestenes, *J. Math. Phys.* **8**, 798 (1967).

[8] D. Hestenes, *J. Math. Phys.* **14**, 893 (1973).

[9] G. Casanova, *C.R. Acad. Sc. Paris*, **266B**, 1551 (1968; **271A**, 817 (1970); **275B**, 53, 121, 267, 399 (1972).

[10] R. Boudet, *C.R. Acad. Sci. Paris*, **272A**, 767 (1971); **278A**, 1063 (1974).

[11] P. Quilichini, *C.R. Acad. Sc. Paris*, **273B**, 1063 (1971).

[12] D. Hestenes, *Found. Phys.* **15**, 63 (1985).

[13] D. Hestenes, *Found. Phys.* **12**, 153 (1982).

APPENDIX: The Kähler Algebra

Since the Dirac-Kähler equation has been a topic for much discussion at this conference, I will provide a dictionary here to show how easy it is to translate the Clifford algebra employed in this paper into Kähler's representation in terms of differential forms. We consider only algebras of spacetime; the generalization to algebras of higher dimension is trivial.

Let Ψ be any element of a Clifford algebra (real or complex) generated by an orthonormal frame of spacetime vectors γ_μ as in the text above. The expansion of Ψ into k-vector parts Ψ_k is given by

$$\Psi = \sum_{k=0}^{4} \Psi_k \, . \tag{A.1}$$

For $k > 0$, the expansion of Ψ_k with respect to a basis can be written

$$\begin{aligned} \Psi_k &= \frac{1}{k!}\, \phi_{\mu_1 \ldots \mu_k} \gamma^{\mu_1} \wedge \ldots \wedge \gamma^{\mu_k} \\ &= \frac{1}{k!}\, \phi^{\mu_1 \ldots \mu_k} \gamma_{\mu_1} \wedge \ldots \wedge \gamma_{\mu_k} \, , \end{aligned} \tag{A.2}$$

where the coefficients may be real or complex, and $\{\gamma^\mu\}$ is the frame reciprocal to $\{\gamma_\mu\}$, as defined by the equation $\gamma^\mu \cdot \gamma_\nu = \delta^\mu_\nu$. Let us write

$$\gamma^\mu \wedge \Psi = \sum_k \gamma^\mu \wedge \Psi_k \, , \qquad \gamma^\mu \cdot \Psi = \sum_k \gamma^\mu \cdot \Psi_k \, , \tag{A.3}$$

so

$$\Box \wedge \Psi = \sum_k \Box \wedge \Psi_k\,, \qquad \Box \cdot \Psi = \sum_k \Box \cdot \Psi_k\,, \tag{A.4}$$

and

$$\Box \Psi = \Box \cdot \Psi + \Box \wedge \Psi \tag{A.5}$$

We can map all this into differential forms by introducing a multivector differential of mixed grade defined by

$$D = \sum_{k=0} D_k\,, \tag{A.6}$$

where $D_0 = 1$ and for $k > 0$,

$$D_k = \mathbf{d}x_1 \wedge \ldots \wedge \mathbf{d}x_k = \frac{1}{k!}\mathbf{d}x^{\mu_1} \ldots \mathbf{d}x^{\mu_k} \gamma_{\mu_1} \wedge \ldots \wedge \gamma_{\mu_k}\,. \tag{A.7}$$

To every k-vector Ψ_k there corresponds a differential k-form Φ_k given by

$$\Phi_k = <\widetilde{D}\Psi_k> = \ \widetilde{D}_k \cdot \Psi_k = \frac{1}{k!}\,\Phi_{\mu_1 \ldots \mu_k} \mathbf{d}x^{\mu_1} \wedge \ldots \wedge \mathbf{d}x^{\mu_k}\,. \tag{A.8}$$

More generally, to every multivector Ψ there corresponds a differential "multiform" Φ given by

$$\Phi = <\widetilde{D}\Psi> = <\Psi\widetilde{D}> = \sum_{k=0}^{4} \Phi_k\,. \tag{A.9}$$

This generalizes the mapping of multivectors into forms which I considered in my first lecture.

Now the *exterior product* of a l-form $\mathbf{d}x^\mu = \gamma^\mu \cdot \mathbf{d}x$ with the form Φ can be defined by

$$\mathbf{d}x^\mu \wedge \Phi = <(\gamma^\mu \wedge \Psi)\widetilde{D}>\,, \tag{A.10}$$

and the *contraction* of Φ with the vector γ^μ can be defined by

$$\gamma^\mu \lrcorner \Phi = <(\gamma^\mu \cdot \Psi)\widetilde{D}>\ . \tag{A.11}$$

Kähler defined a "vee product" for differential forms by writing

$$\mathbf{d}x^\mu \vee \Phi = \mathbf{d}x^\mu \wedge \Phi + \gamma^\mu \lrcorner \Phi\,. \tag{A.12}$$

But (A.10) and (A.11) imply that this is equivalent to

$$\mathbf{d}x^\mu \vee \Phi = <\gamma^\mu \Psi \widetilde{D}>\,, \tag{A.13}$$

which defines the vee product by a linear mapping of the geometric product into forms. Thus, we have established a one-to-one mapping of Clifford algebra onto differential forms. This representation of Clifford algebra by an algebra of differential forms is called the *Kähler algebra*.

The induced mapping of the curl into the differential forms gives us the *exterior derivative*, as expressed by

$$\mathbf{d}\Phi = <\widetilde{D}\Box \wedge \Psi>\ . \tag{A.14}$$

This divergence maps to

$$-\delta\Phi = <\widetilde{D}\Box \cdot \Psi>\ . \tag{A.15}$$

Therefore, the vector derivative maps to

$$(\mathbf{d} - \delta)\Phi = <\widetilde{D}\Box\Psi> = \mathbf{d}x^\mu \vee <\widetilde{D}\partial_\mu \Psi>\ . \tag{A.16}$$

Now the Dirac equation in any of the forms we discussed above can easily be mapped into an equivalent Dirac-Kähler equation in the Kähler algebra. Of course, ideals in the Clifford algebra map into corresponding ideals in the Dirac algebra. I should mention, though, that for ideal spinors on curved manifolds the Kähler derivative defined by (A.16) differs from the derivative in the usual form of the Dirac equation. In particular, it couples minimal left ideals. This point is discussed by Benn and others in their lectures. I want to point out, however, that no such coupling occurs if one employs the operator representation of a Dirac spinor, and there is more than one possible way to define a covariant derivative for spinors. Unfortunately, there appears to be no hope of distinguishing between the various possibilities by any sort of experimental test.