

# THE KINEMATIC ORIGIN OF COMPLEX WAVE FUNCTIONS

David Hestenes

**Abstract.** A reformulation of the Dirac theory reveals that  $i\hbar$  has a geometric meaning relating it to electron spin. This provides the basis for a coherent physical interpretation of the Dirac and Schrödinger theories wherein the complex phase factor  $\exp(-i\varphi/\hbar)$  in the wave function describes electron zitterbewegung, a localized, circular motion generating the electron spin and magnetic moment. Zitterbewegung interactions also generate resonances which may explain quantization, diffraction, and the Pauli principle.

*You know, it would be sufficient to really understand the electron.*  
— Albert Einstein [1]

## 1. Introduction.

Edwin T. Jaynes is one of the great thinkers of twentieth century science [20]. More than anyone else he has deepened and clarified the role of statistical inference in science and engineering. To my mind, his greatest accomplishment has been to recognize that in the evolution of statistical mechanics the principles of physics had gotten confused with principles of statistical inference, and then to show how the two can be cleanly separated to produce a simpler yet more powerful theoretical system.

I share with Ed Jaynes the belief that quantum mechanics suffers from an analogous muddle of probability with physics, which is at the root of the perennial controversy over physical interpretation. Though a Jaynesian revolution of the “quantum muddle” remains elusive, I will report here on a promising possibility that has been overlooked.

One of the puzzling features of quantum mechanics is the fact that “probability amplitudes” are complex numbers whereas probabilities are real. This has inspired a belief that quantum mechanics somehow involves a generalization of the “classical probability concept.” On the contrary, I contend that complex phase factors have a physical origin that has nothing to do with the probability concept *per se*.

The words of Einstein heading this article serve to remind us that the main ideas, as well as the greatest successes, of quantum mechanics have come from studying the electron. To the electron, therefore, we should look for the physical clues needed to resolve the quantum muddle. Now, the Dirac electron theory and its extension to *quantum electrodynamics* is universally recognized as the most well substantiated domain of physics. Strangely, however, it is rarely involved in the discussions of the foundations of quantum mechanics. This is a grievous error, for the *Dirac theory entails an irreducible relation between spin and complex numbers* with undeniable implications for the interpretation of quantum mechanics. Analysis of this relation strongly suggests that the *complex phase factor in the complex*

*function describes a kinematic feature of electron wave motion* and therefore has a physical, rather than statistical, origin.

The argument in support of these contentions has been elaborated elsewhere [3], so we can be satisfied here with an outline of the main ideas. The argument and its implications are developed in three separate stages. Stage I is a reformulation of the Dirac theory which makes its geometric structure explicit, and reveals a connection between spin and the complex imaginary. Stage II assigns a coherent physical interpretation to the Dirac theory which is consistent with its geometric structure. Stage III speculates on the possibility of a deeper theory of electrons to which the Dirac theory is only an approximation.

In the text that follows, major assertions are set off in boxes and supported by just enough discussion to make them intelligible or, at least, stimulating to reader familiar with the Dirac theory.

## 2. STAGE I. Explicating the geometric structure of the Dirac theory.

The crucial first step is to recognize that the set of Dirac matrices  $\gamma_\mu$  can be interpreted as an orthonormal frame of vectors in spacetime, rather than mysterious matrix components of a single vector, as is ordinarily done. This is possible because products of the  $\gamma_\mu$  have well-defined geometric meanings. Thus, the familiar symmetrized product

$$\frac{1}{2}(\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu) = \gamma_\mu \cdot \gamma_\nu = g_{\mu\nu} \quad (1)$$

is nothing other than the *inner product* of vectors defining the metric tensor  $g_{\mu\nu}$ . On the other hand, the antisymmetrized product

$$\frac{1}{2}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) = \gamma_\mu \wedge \gamma_\nu \quad (2)$$

can be identified with the *outer product* of Grassmann algebra, with the  $\gamma_\mu \wedge \gamma_\nu$  composing a basis for the space of *bivectors* (i.e., skew-symmetric tensors of rank two). These two products are thus parts of a single *geometric product*

$$\gamma_\mu\gamma_\nu = \gamma_\mu \cdot \gamma_\nu + \gamma_\mu \wedge \gamma_\nu. \quad (3)$$

Two other geometric entities are generated by this product: the trivectors (or *pseudovectors*)

$$\gamma_\mu \wedge \gamma_\nu \wedge \gamma_\alpha = \epsilon_{\mu\nu\alpha\beta} \gamma^\beta \gamma_5 \quad (4)$$

and the *pseudoscalar*

$$\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3 = \gamma_0 \wedge \gamma_1 \wedge \gamma_2 \wedge \gamma_3. \quad (5)$$

All of these geometric entities are perfectly defined without regarding the  $\gamma_\mu$  as matrices. Every element of the Dirac algebra can be expressed as a linear combination of the 16 basis elements (with  $\mu, \nu = 0, 1, 2, 3$ )

$$I, \quad \gamma_\mu, \quad \gamma_\mu \wedge \gamma_\nu, \quad \gamma_\mu\gamma_5, \quad \gamma_5 \quad (6)$$

where  $I$  is the identity matrix if the  $\gamma_\mu$  are matrices, but  $I = 1$  is the unit scalar if the  $\gamma_\mu$  are regarded as vectors. If the coefficients are real, the matrix algebra generated by

these basis elements is called the *real Dirac algebra*, but the algebra is called the *spacetime algebra* if the  $\gamma_\mu$  are vectors. Thus, it can be concluded that:

The real Dirac algebra is a matrix representation of a generic geometric algebra describing properties of spacetime with no special relation to the quantum mechanics of spin.

The spacetime algebra was first employed in the formulation of electrodynamics and general relativity in [4].

Naturally, it takes some practice to become fully conversant with the spacetime algebra and its geometric significance. In the meantime, readers may continue to regard the  $\gamma_\mu$  as matrices when considering algebraic manipulations.

The next step in explicating the geometric structure of the Dirac theory is to reformulate it in terms of the spacetime algebra. One easy way to do that is to choose a fixed unit spinor  $u$  satisfying the eigenvalue equations

$$\gamma_0 u = u, \tag{7}$$

$$\gamma_2 \gamma_1 u = iu, \tag{8}$$

where  $i$  is the usual unit imaginary of the complex field. Equation (8) is especially significant because it relates  $i$  to the bivector  $\gamma_2 \gamma_1$ , and thus reveals that  $i$  has an implicit geometric meaning in the Dirac theory. This meaning can be made explicit by eliminating  $i$  from the theory in favor of the geometric quantity  $\gamma_2 \gamma_1$ . It is easy to prove that each Dirac spinor  $\Psi$  can be written in the form

$$\Psi = \psi u, \tag{9}$$

where  $\psi$  is a unique *even* element of the *real* Dirac algebra. In other words,  $\psi$  can be written in the form

$$\psi = \alpha_1 + \alpha^{\mu\nu} \gamma_\mu \wedge \gamma_\nu + \alpha_2 \gamma_5, \tag{10}$$

where the eight alphas are real coefficients. Since there is an isomorphism between  $\Psi$  and  $\psi$ , it is fair to refer to  $\psi$  as a spinor or, more specifically, as an *operator representation* of a Dirac spinor.

In terms of  $\psi$  the Dirac equation takes the form

$$\gamma^\mu \left( \partial_\mu \psi \gamma_2 \gamma_1 \hbar - \frac{e}{c} A_\mu \psi \right) = mc \psi \gamma_0 \tag{11}$$

Multiplication by  $u$  on the right shows immediately that this is equivalent to the usual Dirac equation

$$\gamma^\mu \left( i \hbar \partial_\mu - \frac{e}{c} A_\mu \right) \Psi = mc \Psi \tag{12}$$

The significant difference between these isomorphic equations is that (11) involves the real Dirac algebra exclusively. Note how the imaginary factor  $i \hbar$  in (12) has been replaced by the bivector  $\gamma_2 \gamma_1 \hbar$  in (11).

The geometric structure of the Dirac wave function is expressed by the following assertion.

The Dirac wave function has an invariant operator representation

$$\psi = (\rho e^{\beta\gamma_5})^{\frac{1}{2}} R, \quad (13)$$

where  $R$  is a unimodular spinor describing the kinematics of electron motion.

Here unimodularity means that  $R$  can be written in the form

$$R = e^{\frac{1}{2}B}, \quad (14)$$

where  $B$  is a bivector. The relation to kinematics comes from the fact that  $R$  determines a Lorentz transformation of the frame  $\{\gamma_\mu\}$  into a “frame of observables”  $\{e_\mu\}$  given by

$$e_\mu = R\gamma_\mu\tilde{R}. \quad (15)$$

The timelike unit vector  $e_0$  is the direction of the Dirac current, while  $e_3$  is the direction of the “spin,” or polarization, vector. This can be proved by establishing the following relations to standard expressions in the Dirac theory. The components  $J_\mu$  of the *Dirac current*

$$J = \psi\gamma_0\tilde{\psi} = \rho e_0 \quad (16a)$$

are given by

$$J_\mu = J \cdot \gamma_\mu = \tilde{\Psi}\gamma_\mu\Psi. \quad (16b)$$

The components of the *spin vector*

$$s = \frac{1}{2}\hbar e_3 \quad (17a)$$

are given by

$$\rho s_\mu = \frac{1}{2}\hbar\gamma_\mu \cdot (\psi\gamma_3\tilde{\psi}) = \frac{1}{2}\hbar\tilde{\Psi}\gamma_3\gamma_\mu\Psi. \quad (17b)$$

The  $\gamma_\mu$  (or the  $\gamma_0\gamma_\mu$ ) have been interpreted as velocity operators in the Dirac theory. As (16b) shows, this is tantamount to identifying the direction of Dirac current as the local electron velocity, so the  $\gamma_\mu$  are operators only in the trivial sense of picking-out vector components by the inner product with basis vectors.

Spin angular momentum is actually a bivector quantity, though it can be proved from angular momentum conservation in the Dirac theory that the bivector spin  $S$  is related to the spin vector  $s$  in (17a) by

$$S = \gamma_5 s e_0 = \frac{1}{2}\hbar\gamma_5 e_3 e_0. \quad (18a)$$

Moreover,  $\gamma_5 e_3 e_0 = e_2 e_1$ , so

$$S = \frac{1}{2}\hbar e_2 e_1 = \frac{1}{2}R(\gamma_2\gamma_1\hbar)\tilde{R}. \quad (18b)$$

This relates the bivector  $\gamma_2\gamma_1\hbar$  in the Dirac equation (11) to the electron spin. To express it as a relation in the standard Dirac language, multiply (9) by (18b) and use (13) and (8) to get

$$S\Psi = \frac{1}{2}i\hbar\Psi \quad (19)$$

This proves unequivocally that the “imaginary” factor  $i\hbar$  in the Dirac equation is a representation of the electron spin angular momentum  $S$  by its eigenvalue, and the electron wave function is always an “eigenstate” of the spin.

This *fact* that the ubiquitous factor  $i\hbar$  is a representation of electron spin necessarily applies to the Schrödinger equation as well, in as much as it is an approximation to the Dirac equation. It enriches Dirac’s conclusion that *the most fundamental aspect of quantum mechanics is the role of the unit imaginary  $i$*  [5].

### 3. STAGE II. The zbw interpretation.

The identification of the bivector  $S$  as spin rests on the prior identification of

$$\underline{p}_\mu = i\hbar\partial_\mu - \frac{e}{c}A_\mu \quad (20)$$

as *energy-momentum operator*, which is surely one of the fundamental postulates of quantum mechanics. But what shall we make of the “spin” factor  $i\hbar$  in (20)? An answer is suggested by examining the electron kinematics.

In terms of  $\underline{p}_\mu$  the energy-momentum tensor of the Dirac theory is given by

$$T_{\mu\nu} = \tilde{\Psi}\gamma_\mu\underline{p}_\nu\Psi \quad (21)$$

To see what this implies about electron kinematics, note that the derivatives of the kinematics factor  $R$  in (13) can be written in the form

$$\partial_\mu R = \frac{1}{2}\Omega_\mu R, \quad (22)$$

where  $\Omega_\mu$  is a bivector representing the rotational velocity of the frame  $\{e_\nu\}$  under a displacement in the direction  $\gamma_\mu$ .

The local electron momentum  $p_\mu$  is defined by the energy-momentum flux in the direction of the electron velocity, that is,

$$\rho p_\mu = (e_0 \cdot \gamma^\nu)T_{\mu\nu}, \quad (23)$$

where  $\rho$  is the probability density defined by (13). Evaluated with (21) and (22), (23) yields

$$p_\mu = S \cdot \Omega_\mu - \frac{e}{c}A_\mu \quad (24)$$

The term  $S \cdot \Omega_\mu$  has precisely the form of a rotational kinetic energy. Thus, the intrinsic energy (mass) of the electron is associated with rotational motion in the spacelike plane of the spin  $S$ .

To give this fact a literal physical interpretation wherein the spin  $S$  is generated by electron motion, the velocity of the electron must be redefined to maintain a component

in the spin plane. Accordingly, the most natural definition of electron velocity is the *null vector*

$$u = e_0 - e_2. \quad (25)$$

This leads to a *self-consistent interpretation of the Dirac theory with the following features*:

- (a) The electron is modeled as a structureless point particle travelling at the speed of light along a helical lightlike trajectory in spacetime.
- (b) The helical trajectory has a diameter on the order of a Compton wavelength, and a circular frequency on the order of twice the de Broglie frequency  $mc^2/\hbar \approx 10^{21}\text{s}^{-1}$ .
- (c) The helical motion generates electron spin and may be attributed to magnetic self-interaction.
- (d) Each solution of the Dirac equation determines an infinite family of such helices and a probability distribution for the electron to be found on any given helix.
- (e) The center of curvature for each helix lies on a streamline of the Dirac current.

All these assertions are consistent with the mathematical form of the Dirac theory, and they supply the mathematical structure of the theory with the most complete and coherent physical interpretation available. They constitute a generalization of the *zitterbewegung interpretation* originally proposed by Schrödinger.

The *zitterbewegung* (zbw) is reflected in the structure of the kinematic factor  $R$  of the Dirac wave function (13) by writing it in the form

$$R = R_0 e^{-\gamma_2 \gamma_1 \varphi / \hbar}. \quad (26)$$

The helical motion of the electron can be visualized as a particle moving in a circle lying in the spacelike plane of the spin  $S$  while the center of the circle is translated along a streamline of the Dirac current. The *phase angle*  $\varphi$  in (26) represents the angular displacement on the circle (in units of angular momentum), lying in  $\gamma_2 \gamma_1$ -plane, while  $R_0$  represents a Lorentz transformation which rotates this plane into the spin plane. Inserting (26) into (22) and using (18b), one finds

$$\Omega_\mu = 2(\partial_\mu R)\tilde{R} = 2(\partial_\mu R)\tilde{R}_0 + (\partial_\mu \varphi)S^{-1}. \quad (27)$$

The term involving  $R_0$  describes relativistic effects as the electron is accelerated by external forces, and it is always smaller than the last term. By keeping only the dominant term, then, insertion of (27) into (24) yields

$$p_\mu = \partial_\mu \varphi - \frac{e}{c} A_\mu, \quad (28)$$

which shows explicitly that the electron energy-momentum can be attributed to the *circular zitterbewegung*. In the same approximation, we can set  $R_0 = 1$  and the Dirac wave function reduces to

$$\psi_0 = \rho^{\frac{1}{2}} e^{-\gamma_2 \gamma_1 \varphi / \hbar} \quad (29)$$

This is exactly the form of the Schrödinger wave function. Thus, we conclude that

The complex phase factor  $\exp(-i\varphi/\hbar)$  in both the Dirac and Schrödinger wave functions describes kinematics of electron motion, specifically the circular zitterbewegung.

Schrödinger theory describes the dominant component of the zitterbewegung. That's why it is such a successful approximation to the Dirac theory.

Of course, the challenge to this interpretation is to devise experimental tests which show that the helical zitterbewegung is a real physical phenomenon. One such test is a prediction of deviations from the Mott-scattering cross section when the impact parameter is on the order of a Compton wavelength [6].

#### 4. STAGE III. Zitterbewegung Interactions.

If the circular zitterbewegung is a real physical phenomenon and not merely a picturesque metaphor, then we should expect its presence to be manifest in the electron's electromagnetic field. Indeed, to explain the electron's static magnetic dipole field as a consequence of the zitterbewegung, we must regard the electron as a point charge for which the average motion over a zitterbewegung period is an effective current loop. But the same assumption implies that the zitterbewegung must generate high-frequency fluctuations about that average; call this the *zitterbewegung field* of the electron. The frequency of these fluctuations ( $\sim 10^{21}\text{s}^{-1}$ ) is too high to observe directly, but it has been suggested that zitterbewegung fields are responsible for some of the most peculiar features of quantum mechanics [7].

The possibility of zbw interactions is not contemplated in the Dirac theory, though to some degree it may be inherent in the theory and its extension to quantum electrodynamics. In any case, it appears that a complete mathematical treatment of zbw interactions requires new physical assumptions which have not yet been formulated and analyzed, so the best that can be done at this time is a qualitative analysis of possible consequences.

We begin with the picture of each electron as the seat of a bound electromagnetic field fluctuating with the electron's zbw frequency, and we set aside the question of what theoretical assumptions are necessary to make this possible. Every electron, of course, is perturbed by the zbw fields of other electrons, with doppler-shifted frequencies due to their motions. Conceivably, this random background of electromagnetic fluctuations can play the role of the vacuum field in quantum electrodynamics, and a stochastic term in the electron equations of motion is needed to account for its effect. In that case, the Casimir effect could be explained by adapting well-known arguments.

A more significant and prominent feature of zbw interactions is the likelihood of resonances. Three kinds of resonances are of special interest:

(1) *Electron diffraction* is usually “explained” by invoking an interference metaphor. However, an explanation in terms of “quantized momentum exchange” is equally consistent with the formalism of quantum mechanics as well as a strict particle interpretation. Moreover, the zbw field provides a physical mechanism for this exchange. When an electron is incident on a crystal, it is preceded by its zbw field which is reflected from the crystal back to the electron. The conditions for *resonant momentum exchange* are met when the reflected field is resonant with the electron’s zbw, which must occur at the Bragg angles if the explanation is correct.

(2) *Quantized energy states* appear when the frequency of an electron’s orbital motion is resonant with a harmonic of its zbw frequency, as is implied by a zbw interpretation of solutions to Schrödinger’s equations. As in diffraction, the underlying causal mechanism may be a resonance of the electron’s zbw with the reflection of its own zbw field off the atomic nucleus.

(3) *The Pauli principle* may be explained as a zbw resonance between two electrons, mediated by their zbw fields. Such a resonance can occur only when the electrons are in the same state of translational motion, as the Pauli principle requires.

Other quantum phenomenon, such as barrier penetration and the Lamb shift, can also be explained qualitatively by the zbw. The problem remains to make these explanations fully quantitative.

## References

- [1] This quotation was recollected and explained by Valentine Bargmann. In H. Woolf (ed.), *Some Strangeness in Proportion*, a Centennial symposium to celebrate the achievements of Albert Einstein. Addison-Wesley, Reading, MA (1980).
- [2] D. Hestenes, “E.T.Jaynes: Papers on Probability, Statistics and Statistical Physics,” *Found. Phys.*, **14** 187–191 (1984).
- [3] D. Hestenes, “The Zitterbewegung Interpretation of Quantum Mechanics,” *Found. Phys.*, **20**, 1213–1232 (1990). Further references are contained therein.
- [4] D. Hestenes, *Space-Time Algebra*, Gordon and Breach. New York (1966, reprinted with corrections 1992).
- [5] C.N. Yang, “Square root of minus one, complex phases and Erwin Schrödinger.” In C.W. Kilminster (ed.), *Schrödinger*, Cambridge (1987), p. 53–64.
- [6] M.H. MacGregor, “KeV Channeling Effects in the Mott Scattering of Electrons and Positrons,” *Found. Phys. Letters*, **5**, 15–24 (1992).
- [7] D. Hestenes, “Quantum Mechanics from Self-Interaction,” *Found. Phys.* **15**, 63–87 (1985).