

THE THEORY
OF
PARTICLES OF SPIN 1/2
(DIRAC ELECTRONS)

BY

Louis de BROGLIE
PERPETUAL SECRETARY OF THE ACADEMY OF SCIENCES

Translated by D. H. Delphenich

PARIS
GAUTHIER-VILLARS, PRINTER-EDITOR
LIBRARY OF THE BUREAU OF LONGITUDES OF L'ÉCOLE POLYTECHNIQUE
Quai des Grands-Augustins, 55

—
1952

PREFACE

In the present volume, I have presented the theory of particles of spin $1/2$ (i.e., Dirac electrons) by comparing my viewpoint on some great problems to the ones that other authors have pointed out in recent papers.

I begin by recalling the general principles of wave mechanics and its physical interpretation. I then introduce the notion of spin of a particle and examine its various aspects. I then present the theory of the electron, when it is considered to be a corpuscle of spin $1/2$ (i.e., Dirac's theory).

I shall not insist here upon phenomena that have received satisfactory explanations, even though they resisted any complete explanation by the old theories. For those questions, I shall refer to my book *L'Électron Magnétique* ⁽¹⁾.

By contrast, I have analyzed the relativistic dynamics of spinning fluids and spinning particles that is due to Weyssenhoff in order to show the link between it and some concepts that were presented before.

The last part of this work is dedicated to the possibility of measuring the spin of the electron. The validity of Bohr's arguments that tend to show that it is impossible to measure the electron spin directly seem to me to be generally limited to case in which the velocities are small in comparison to that of light. Finally, Pauli's opinion that the point mechanics of a Dirac electron is identical to the point mechanics of an electron without spin contradicts my conclusions. With the present state of research, an even deeper examination will not permit us to say more about that point.

The presentation that was given to this book was intentionally chosen in order to preserve the character of reality to those reflections.

I would like to warmly thank Michel Cazin for the very important assistance that he gave me in the publication of the present work.

Louis de Broglie

⁽¹⁾ *L'Électron Magnétique*, Hermann, Paris, 1934.

TABLE OF CONTENTS

	Page
CHAPTER I	
NON-RELATIVISTIC WAVE MECHANICS WITH ONE WAVE FUNCTION.....	1
1. General ideas and equations of wave mechanics.....	1
2. Wave equations of wave mechanics.....	4
3. New conception of the quantities attached to a corpuscle.....	6
4. Proper values and proper functions of a linear, Hermitian operator.....	8
5. Continuous spectrum of the Hamiltonian operator for a free corpuscle...	10
CHAPTER II	
PHYSICAL INTERPRETATION OF WAVE MECHANICS.....	13
1. General principles.....	13
2. Algebraic matrices and their properties.....	16
3. Operators and matrices in wave mechanics.....	18
4. Mean values and field quantities in wave mechanics.....	19
5. First integrals in wave mechanics.....	21
6. Precise form for the uncertainty relations.....	22
CHAPTER III	
QUANTUM THEORY OF KINETIC MOMENTS AND SPINS.....	25
1. Orbital kinetic moment.....	25
2. Kinetic moment and the group of spatial rotations.....	27
3. General results relating to the proper values of operators that satisfy the non-commutation relations (III.a).....	29
4. Spin.....	31
CHAPTER IV	
PROPER KINETIC MOMENTS FROM THE RELATIVISTIC VIEWPOINT...	34
1. Generalities.....	34
2. Relativistic representation of orbital kinetic moment.....	35
3. Study of proper kinetic moment from the relativistic viewpoint.....	37
4. General relativistic theory of proper kinetic moments.....	42
5. Relativistic aspect of proper magnetic moments.....	46
6. Relationship between the magneto-electric moment and spin.....	48
7. Jan van Weyssenhoff's theory.....	49

	Page
CHAPTER V	
DIRAC'S THEORY OF THE SPINNING ELECTRON.....	57
1. The wave equations of the spinning electron.....	57
2. The relativistic invariance of the Dirac equations.....	62
3. The spin of the electron in Dirac's theory.....	66
CHAPTER VI	
FORMALISM AND PHYSICAL INTERPRETATION OF DIRAC'S THEORY.....	71
1. General formalism of relativistic wave mechanics for the Dirac electron	71
2. The field quantities that are defined by Dirac's theory.....	73
3. Monochromatic plane waves in Dirac's theory.....	76
4. The density-current quadri-vector and its decomposition.....	80
CHAPTER VII	
NEGATIVE-ENERGY SOLUTIONS IN DIRAC'S THEORY.....	86
1. The negative-energy plane wave.....	86
2. Incomplete character of the system of positive-energy waves.....	89
3. Dirac's theory of "holes".....	93
CHAPTER VIII	
SCHRÖDINGER'S <i>ZITTERBEWEGUNG</i>	97
1. The center of gravity of probability in wave mechanics.....	97
2. Ehrenfest's theorem.....	98
3. Schrödinger's <i>Zitterbewegung</i>	100
CHAPTER IX	
POSSIBILITY OF MEASURING ELECTRON SPIN.....	104
1. Present thinking about the question.....	104
2. Action of a magnetic field upon the proper magnetic moment.....	104
3. Measuring the magnetic field that is produced by an electron.....	106
4. Compensating for the Lorentz force with an electric field.....	107
5. Stopping an oriented electron by a magnetic field gradient.....	110
6. Measuring the proper kinetic moment (spin).....	111
7. Conclusion.....	112

CHAPTER X	Page
PASSING TO THE GEOMETRICAL OPTICS APPROXIMATION IN RELATIVISTIC WAVE MECHANICS.....	113
1. The geometric optics approximation.....	113
2. The WKB method of successive approximations.....	116
3. Passing to geometrical optics in Dirac's theory.....	117
4. The WKB method in Dirac's theory.....	123
5. Applying Pauli's line of reasoning when one starts with Weysenhoff's point mechanics.....	127
6. Calculating the functions C_1 and C_2 in the case of the motion of an electron moving in the direction of a permanent, homogeneous, magnetic field.....	130
7. Return to the passage to geometrical optics in Dirac's theory.....	134
8. General proof of formulas [X.i] and [X.j] upon starting from the Gordon decomposition.....	144
9. Interpretation of the condition $F(a_k^{(0)}) = 0$	147
10. Verification of the relation $(g_{(d_0)}^4)_0 = 0$ in two simple special cases.....	149
11. Conclusion. Numerical values.....	151

CHAPTER I

NON-RELATIVISTIC WAVE MECHANICS WITH ONE WAVE FUNCTION

1. GENERAL IDEAS AND EQUATIONS OF WAVE MECHANICS

The idea that has served as the point of departure for wave mechanics has been the following one: Since there exists a corpuscular aspect, as well as a wave aspect, to light that are coupled by the relation:

$$\text{energy} = h \times \text{frequency},$$

in which Planck's quantum constant h appears, it is natural to assume that there exists a corpuscular aspect and wave-like aspect to matter, as well, and the latter has been known to me for some time. Those two aspects must be coupled by some general relations in which Planck's constant figures that must be treated as special cases of the relations that relate to light.

In order to develop that idea, one must seek to associate a periodic element with the concept of corpuscle. Imagine a corpuscle that moves with a uniform, rectilinear motion in a certain direction in the absence of any field. We fix our attention on just the state of motion of the corpuscle, which is an abstraction that is made from its position in space. That motion is performed in a certain direction that we take to be the z -axis, and it is defined by the two quantities of "energy" and "quantity of motion," whose relativistic expressions as functions of the proper mass m_0 of the corpuscle and its velocity $v = \beta c$ are given by the formulas:

$$W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}}, \quad \mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \beta^2}}, \text{ resp.,}$$

so one can infer the relation:

$$|\mathbf{p}| = p = \frac{Wv}{c^2}.$$

The state of motion is then found to be defined in a certain Galilean reference system for an observer A that employs the coordinates x, y, z, t .

Now, let there be another observer B that possesses a velocity of \mathbf{v} in the direction Oz with respect to the first one; in other words, it is coupled to the motion of a corpuscle. We can suppose that B has chosen an axis $O_0 z_0$ that slides along Oz and axes $O_0 x_0$ and $O_0 y_0$ that are parallel to Ox and Oy , resp. That being the case, the coordinates x_0, y_0, z_0, t_0 that are employed by B will be linked with the coordinates x, y, z, t of the observer A by the Lorentz transformation formulas:

$$x_0 = x, \quad y_0 = y, \quad z_0 = \frac{z - vt}{\sqrt{1 - \beta^2}}, \quad t_0 = \frac{t - \beta z / c}{\sqrt{1 - \beta^2}}.$$

Now, the velocity of the corpuscle is zero for observer B : He will then set the values of the energy and the quantity of motion equal to:

$$W_0 = m_0 c^2, \quad \mathbf{p}_0 = 0, \text{ resp.}$$

According to our basic idea, we must now seek to introduce a periodic element, and we attempt to define the desired periodic element in the form of a wave that is stationary in the proper system of the corpuscle (system of the observer B). We then set:

$$\psi_0 = A e^{2\pi i \nu_0 t_0}$$

and suppose that A is constant: ν_0 is the *proper* frequency of the wave, which must depend upon the nature of the corpuscle envisioned. What value must we give to that constant? We must obviously seek to define it by starting with a non-zero value that is characterized by the corpuscle in its proper system, and we have the energy W_0 at our disposal as just such a quantity. Being given the role that is played by the quantum constant in all of quantum theory, it would be natural to set:

$$\nu_0 = \frac{W_0}{h} = \frac{m_0 c^2}{h},$$

which is analogous to Einstein's relation for photons.

How does the periodic element that we just defined for the observer B manifest itself for the observer A ? Upon supposing the simplest hypothesis that ψ is a relativistic invariant, in order to obtain the expression for the wave for A , it will suffice to substitute the fourth equation for the Lorentz transformation in its expression for B :

$$\psi(x, y, z, t) = A e^{2\pi i \nu(t - z/v)},$$

with:

$$\nu = \frac{\nu_0}{\sqrt{1 - \beta^2}}, \quad V = \frac{c}{\beta} = \frac{c^2}{v}.$$

Hence, for the observer A , the phases of the periodic element that was introduced are distributed like the phases of a monochromatic plane wave whose frequency ν and speed of propagation V of the phase have the indicated values.

Upon comparing the preceding equations, one will find that $W = h\nu$, which is a relation that will obviously be valid in all Galilean systems, since nothing distinguishes the observer A from the other Galilean observers. For the wave length of the wave ψ , one will find, from the usual definition, that:

$$\lambda = \frac{V}{\nu} = \frac{c^2}{\nu} \cdot \frac{h}{W} = \frac{h}{p},$$

which is a fundamental formula that takes the approximate form:

$$\lambda = \frac{h}{mc}$$

for small velocities, and that formula has been verified with great precision by the experiments on the diffraction of electrons and other particles (including neutrons) by crystals, as well as Börsch's experiment on the diffraction of electrons by the edge of a screen.

For a particle whose velocity is very close c , one will find that:

$$\nu = V = c, \quad W = h\nu, \quad p = \frac{h}{\lambda} = \frac{h\nu}{c}.$$

One will then find the fundamental formulas of the theory of photons (viz., Einstein's light quanta).

We can now write the form of ψ in the system A :

$$\psi = A e^{\frac{2\pi i}{h}(Wt - p_z z)},$$

and more generally, if the rectangular axes have an arbitrary orientation:

$$\psi(x, y, z, t) = A e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} = A e^{\frac{2\pi i}{h}(Wt - \mathbf{p} \cdot \mathbf{r})}.$$

One then sees that, up to the factor $2\pi / h$, the phase of the wave is equal to the Hamiltonian action of the corpuscle. Upon asserting that proportionality between the Hamiltonian action of the corpuscle and the phase of the wave that it is associated with, one will perceive that the principle of stationary action that is valid for the dynamics of corpuscles is only a translation of the Fermat's principle that is valid for the associated wave. However, wave theory teaches us that Fermat's principle is valid only in the domain in which geometrical optics is useful, and will lose its value in the domain of physical optics, properly speaking. One will then arrive at the fundamental idea that the old mechanics (also in its classical Newtonian form, as well as in its relativistic form) is only an approximation that has the same domain of validity as the geometrical optics approximation. From that, one is led to imagine that necessity of constructing a new mechanics that will relate to the old mechanics in the same way that wave optics relates to geometrical optics.

2. WAVE EQUATIONS OF WAVE MECHANICS

We have arrived at the idea that one must associate a corpuscle with a wave that is represented by a function $\psi(x, y, z, t)$ that will generally be non-zero in an extended region of space. In other words, we combine the idea of corpuscle with that of *field*, in the sense of a physical field – namely, the field ψ .

The wave function ψ must satisfy a certain “equation of propagation” that will replace the classical Newton equations and serve as the basis for the new mechanics. We shall seek to write that equation without being preoccupied with satisfying the demands of relativity theory, for the moment. We will then obtain a non-relativistic wave mechanics that will be valid only for motions with velocities that are much less than c .

Consider a corpuscle of mass m that displaces in a force field that is derived from a potential function $U(x, y, z, t)$. Let \mathbf{p} be the impulse of the corpuscle, and let E be its total energy:

$$E = \frac{1}{2}mv^2 + U(x, y, z, t) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z, t).$$

By definition, we call a function of the coordinates, the components of impulse, and time that gives the value of energy a *Hamiltonian function* $H(x, y, z, t, p_x, p_y, p_z)$. Here, one will have:

$$H(x, y, z, t, p_x, p_y, p_z) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z, t).$$

The development of wave mechanics has shown that one will get the equation of propagation for waves ψ upon starting from the Hamiltonian function H by the following process: One begins by replacing each of the momenta p_k in the expression for H with the operator $-\frac{h}{2\pi i} \frac{\partial}{\partial x_k}$, which will yield the operator:

$$H_{\text{op}} = H\left(x, y, z, t, -\frac{h}{2\pi i} \frac{\partial}{\partial x}, -\frac{h}{2\pi i} \frac{\partial}{\partial y}, -\frac{h}{2\pi i} \frac{\partial}{\partial z}\right),$$

which is called the “Hamiltonian operator,” or more briefly, the *Hamiltonian*. One then writes:

$$\frac{h}{2\pi i} \frac{\partial}{\partial t} \psi = H_{\text{op}} \psi,$$

in which ψ is the wave function. One will then obtain the wave equation for the corpuscle considered.

Upon specifying the form of H_{op} , one will get:

$$(I.a) \quad \frac{h}{2\pi i} \frac{\partial}{\partial t} \psi = \frac{-h^2}{8\pi^2 m} \Delta \psi + U \psi,$$

or further:

$$\Delta \psi - \frac{8\pi^2 m}{h^2} U \psi = \frac{4\pi i m}{h} \frac{\partial \psi}{\partial t}.$$

Since that equation of propagation has first order in time, it will, in principle, permit one to calculate the form of the function ψ at any instant t when one knows its form $\psi(x, y, z, t_0)$ at the initial instant t_0 .

Equation (I.a) has complex coefficients: The function ψ has an essentially complex character. We let F^* denote the quantity that is the complex conjugate of the complex quantity F . The equation that is satisfied by the function ψ^* is the equation that is the complex conjugate of (I.a).

If we set, by definition:

$$\rho = \psi \psi^*, \quad \mathbf{f} = \frac{h}{4\pi i m} (\psi^* \text{grad } \psi - \psi \text{grad } \psi^*)$$

then we can show that upon starting from the equations that were satisfied by ψ and ψ^* , we will have:

$$\frac{\partial \rho}{\partial t} + \text{div } \mathbf{f} = 0.$$

That equation has the classical form of an equation of continuity. If the wave ψ occupies a domain D (finite or infinite) and is zero on the boundary of D then one will infer from the preceding equation that:

$$\int_D \rho d\tau = \int_D |\psi|^2 d\tau$$

is constant in the course of time. Like the function ψ , the solution of a *linear* equation is defined only up to a multiplicative constant, so one can choose that constant in such a fashion that one will *continually* have:

$$\int_D |\psi|^2 d\tau = 1.$$

One then says that ψ is *normalized*, and we assume that all functions must always be normalized, which is a hypothesis that justifies the physical interpretation of the quantity $|\psi|^2$ that will be given later on. Even when it has been normalized, the function ψ will further contain an arbitrary factor $e^{i\alpha}$ of norm 1.

One shows that wave mechanics, whose basic equation we just obtained, will admit the old mechanics as an approximation to the same degree of approximation as geometrical optics.

A particularly important case is the one in which the function U does not depend upon time (e.g., a permanent exterior field). The wave equation for ψ will then admit

“monochromatic” solutions that contain time only by way of an exponential factor of the form $e^{\frac{2\pi i}{h}Et}$. Such a wave will be a solution of the equation:

$$\Delta\psi + \frac{8\pi^2 m}{h^2} [E - U(x, y, z, t)] \psi = 0,$$

which would result by substitution in equation (I.a).

In the even-more-special case in which $U = 0$, one will find solutions of the wave equation that have the “monochromatic plane wave” type:

$$\psi = Ae^{\frac{2\pi i}{h}[Et - \sqrt{2mE}(\alpha x + \beta y + \gamma z)]},$$

in which A is a constant, and α, β, γ are the direction cosines of its direction, which are coupled by $\alpha^2 + \beta^2 + \gamma^2 = 1$. That solution represents a monochromatic plane wave of frequency $\nu = h / c$ and wave length:

$$\lambda = \frac{h}{\sqrt{2mE}} = \frac{h}{p} = \frac{h}{mv}$$

that propagates in the direction of α, β, γ . We will then recover (with only an unimportant difference in the definition of the frequency ν) the monochromatic plane wave that, from its inception, has made a uniform, rectilinear motion that corresponds to a corpuscle of mass m , energy E , and quantity of motion mv in the direction α, β, γ in the absence of a field.

3. NEW CONCEPTION OF THE QUANTITIES ATTACHED TO A CORPUSCLE

In the method that we just developed, one replaces the quantities p_x, p_y, p_z that represented the quantity of motion of a corpuscle in the old mechanics with the operators $-\frac{h}{2\pi i} \frac{\partial}{\partial x}, -\frac{h}{2\pi i} \frac{\partial}{\partial y}, -\frac{h}{2\pi i} \frac{\partial}{\partial z}$. That idea of substituting or making “operators” correspond to classical “quantities” has been raised to the status of a general principal in the evolution of wave mechanics. One assumes that any measurable quantity (i.e., observable) that is defined in classical mechanics must correspond to a certain operator in the new mechanics. In order to form that operator upon starting from the classical expression for the quantity, when it has been expressed in terms of the Lagrange variables x, y, z, p_x, p_y, p_z , one will be led to adopt the following rule: As we already know, we make the quantities p_x, p_y, p_z correspond to the operators $-\frac{h}{2\pi i} \frac{\partial}{\partial x}, -\frac{h}{2\pi i} \frac{\partial}{\partial y}, -\frac{h}{2\pi i} \frac{\partial}{\partial z}$, resp. We make the quantities x, y, z correspond to the operators $x \cdot, y \cdot, z \cdot,$

resp.; i.e., multiplication by x , y , z , resp. We will only have to replace the canonical variables in the classical expression for the quantity in question as a function of x , y , z , p_x , p_y , p_z with the corresponding operators in order to get the operator that corresponds to that quantity. That operator can contain time as a numerical parameter if the expression for the quantity contains it. We point out that it is precisely upon applying the preceding method to the quantity of “energy” that we obtained the Hamiltonian operator.

Upon applying that method to the component along the z -axis of the kinetic moment of a corpuscle with respect to the origin, one will find the operator:

$$(M_x)_{\text{op}} = (x p_y - y p_x)_{\text{op}} = -\frac{h}{2\pi i} \left(x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} \right) = -\frac{h}{2\pi i} \frac{\partial}{\partial \varphi},$$

in which φ is the azimuth, when measured around Oz .

The operators that one then forms in wave mechanics in order to make them correspond to measurable quantities are generally complex operators that belong to the Hermitian operators that we shall now define: We say that an operator is *Hermitian* in a domain D if:

$$\int_D f^* A g d\tau = \int_D g^* A f d\tau,$$

whenever f and g are two arbitrary functions in D that are subject to only the condition that they must be finite, uniform, and continuous in that domain, and zero on the boundary of D , in such a fashion that the surface integrals that appear under integration by parts of \int_D will be zero.

All of the operators in wave mechanics that correspond to observable quantities are Hermitian. One can, for example, verify that easily for H and M , as they were defined above. We shall see the physical significance of that fact later on.

The operators of the wave mechanics are not just Hermitian, they are also linear; i.e.:

$$A (\varphi_1 + \varphi_2) = A \varphi_1 + A \varphi_2, \quad A (c \varphi) = c A \varphi,$$

in which c is an arbitrary complex constant.

There is also good reason to distinguish two categories of operators in wave mechanics: The “complete” operators that address all of the variables in the domain D (here: x , y , z) and the “incomplete” operators, which address only some of those variables. The operator H has the type of a complete operator, while M_z is incomplete.

Briefly, in wave mechanics, one makes any physically-observable quantity that is attached to a corpuscle correspond to linear, Hermitian operator that is generally complex. However, it is quite obvious that when one performs a precise measurement of that quantity, that measurement will be expressed by a *real* number. Wave mechanics must then be able to say what the real numbers are that precise measurements can give us for the values of a certain physical quantity.

We must then be able to deduce a series of real numbers that represent all of the possible results of the measurement of a certain measurable quantity from the linear, Hermitian operator that corresponds to it in the new mechanics. Now, that is possible

precisely because linear, Hermitian operators like the ones that are employed by wave mechanics have a sequence of “proper values” that are always real numbers. We shall study that point.

4. PROPER VALUES AND PROPER FUNCTIONS OF A LINEAR, HERMITIAN OPERATOR

Let A be a linear, Hermitian operator. Write the equation:

$$(I.b) \quad A\varphi = \alpha\varphi,$$

in which φ is a function of x, y, z , and α is a constant. Time t can figure as a parameter in A, φ , and α . By definition, we say the *proper values of the operator A in the domain D* to mean the values of the constant α for which the preceding equation will have at least one solution, which we call a *proper function*, and which enjoys the following properties: It is uniform and continuous in D , and the integral of the square of its modulus in D is convergent. If D is infinite then the latter condition will imply that φ must decrease very rapidly at infinity in order to assure the aforementioned convergence. Moreover, if D is finite then φ must be zero on the boundary of D .

We assume the existence of proper values for the operators that are encountered in wave mechanics, and we shall show that they are real. Indeed, one infers from (I.b) and its conjugate that:

$$\int_D [\varphi^* A\varphi - \varphi A^* \varphi^*] d\tau = (\alpha - \alpha^*) \int_D |\varphi|^2 d\tau.$$

Since A is Hermitian, by hypothesis, the first number will be zero: The integral of the second number is essentially positive, so one must have $\alpha = \alpha^*$, and α must be real.

The set of proper values of (I.b) defines the *spectrum* of that equation. If those proper values are isolated then the spectrum will be discontinuous; i.e., it will be a “line spectrum.” On the contrary, if the sequence of proper values is continuous then it will be a “continuous spectrum.” A spectrum can be partially continuous and partially discontinuous, moreover.

We first address the discontinuous spectrum. Let α_i denote an isolated proper value; there will then exist at least one proper function $\varphi_i(x, y, z, t)$ that corresponds to it. We show that the set of proper functions with a discontinuous spectrum will define an *orthogonal set*; i.e., if φ_i and φ_j are two proper functions that correspond to distinct proper values α_i and α_j , resp, then we will have:

$$\int_D \varphi_i^* \varphi_j d\tau = 0.$$

Indeed, since all of the α_i are real, we can infer from (I.b) and its conjugate that:

$$\int_D [\varphi_i^* A\varphi_j - \varphi_j A^* \varphi_i^*] d\tau = (\alpha_i - \alpha_j) \int_D \varphi_i^* \varphi_j d\tau,$$

and since the left-hand side is zero (since A is Hermitian) and $\alpha_i \neq \alpha_j$, the stated result will ensue.

The preceding proof will break down for two proper functions that correspond to the same proper value. When that case presents itself, one says that one is dealing with a *multiple* or *degenerate* proper value. Let α_i be such a proper value that corresponds to p linearly-independent proper functions φ_{i1} , φ_{i2} , ..., φ_{ip} . If one knows those p independent proper functions then one can replace them with p linearly-independent linear combinations of the φ_{i1} , φ_{i2} , ..., φ_{ip} , because since equation (I.b) is linear, such combinations will again be solutions for the same value of α_i or α . One easily sees that one can choose those linear combinations in such a fashion that they are orthogonal. Even in that case, one can then choose the proper functions in such a fashion that one will have an orthogonal system.

Since proper functions are obviously defined only up to a complex multiplicative constant, due to the linear character of (I.b), one can choose the modulus of that constant in such a fashion that:

$$\int_D |\varphi_i|^2 d\tau = 1.$$

The function φ_i is therefore “normalized”: viz., it also contains an arbitrary factor $e^{i\alpha}$ of modulus unity. Once normalized, the functions φ_i will define an *orthogonal system* such that:

$$\int_D \varphi_i^* \varphi_j d\tau = \delta_{ij},$$

in which δ_{ij} is the Kronecker symbol ($\delta_{ij} = 0$ if $i \neq j$; $\delta_{ij} = 1$).

We pass on to the case of a continuous spectrum. If A possesses a continuous spectrum then any proper value in that spectrum will correspond to a proper function $\varphi(x, y, z, \alpha)$, in which we write α as a variable, because it can vary continuously throughout the spectrum. One easily proves that any proper function of the continuous spectrum is orthogonal to the proper functions of the discontinuous spectrum (if there is one). However, in order to express the idea that the proper functions of the continuous spectrum are normalized and mutually-orthogonal, if one is to avoid certain problems with convergence then it can be useful to consider, instead of the proper functions $\varphi(x, y, z, \alpha)$ themselves, the expressions:

$$\int_{\alpha}^{\alpha+\Delta\alpha} \varphi(x, y, z, \alpha) d\alpha,$$

which are called “proper differentials” and correspond to intervals $(\alpha, \alpha + \Delta\alpha)$ that are as small as one desires in the domain of variation for the parameter α . Physically, that substitution will correspond to the one that one carries out in the old wave theory when one considers a group of waves that is composed of the superposition of some waves with very close frequencies instead of a monochromatic plane wave, which is an abstraction. In order to express the orthogonality of the proper differentials, one must replace the relations that are valid for the continuous spectrum with the following one:

$$\frac{1}{\Delta\alpha} \int_D d\tau \left[\int_{\alpha'}^{\alpha'+\Delta\alpha'} \varphi(x, y, z, \alpha) d\alpha \right] \left[\int_{\alpha''}^{\alpha''+\Delta\alpha''} \varphi(x, y, z, \alpha) d\alpha \right]^* = \delta_{\alpha'\alpha''}.$$

The proper functions of complete operators in wave mechanics possess the important property that they define a *complete system*. That amounts to saying that under very broad conditions, a function that is defined in the domain D of the variables that bear upon an operator A can be developed into a sum of proper functions of that operator. For example, if $f(x, y, z)$ is a function of the variables x, y, z then it can be very generally developed in proper functions of a complete Hermitian operator in the form:

$$f(x, y, z) = \sum_i c_i \varphi_i(x, y, z) + \int c(\alpha) \varphi_i(x, y, z, \alpha) d\alpha,$$

in which the sum \sum relates to the discontinuous spectrum, and the integral, to the continuous one. In the preceding development, we can exhibit the proper differentials of the continuous spectrum by writing:

$$f(x, y, z) = \sum_i c_i \varphi_i(x, y, z) + \sum_{\Delta\alpha} c(\alpha) \int_{\alpha}^{\alpha+\Delta\alpha} \varphi_i(x, y, z, \alpha) d\alpha.$$

(More rigorously, one must introduce the notion of “convergence in the mean” here, which we shall not do.)

Upon using the formulas that express the orthogonal character of proper functions of the discontinuous spectrum and the proper differentials of the continuous spectrum, we will find that:

$$c_i = \int_D \varphi_i^* f(x, y, z) d\tau, \quad c(\alpha) = \frac{1}{\Delta\alpha} \int_D d\tau \left[\int_{\alpha}^{\alpha+\Delta\alpha} \varphi(x, y, z, \alpha) d\alpha \right]^* f(x, y, z).$$

The coefficients c_i and $c(\alpha)$ are often called the *Fourier coefficients* of the development of the function $f(x, y, z)$ in proper functions of the operator A . The Fourier series and integral are simple examples of that type of development. Time can figure as a numerical parameter in the expression for the c_i and $c(\alpha)$.

5. CONTINUOUS SPECTRUM OF THE HAMILTONIAN OPERATOR FOR A FREE CORPUSCLE

The proper value equation for the Hamiltonian operator can be written:

$$H\varphi = E\varphi,$$

in which E is the constant α of the general case. For a free corpuscle:

$$U = 0 \quad \text{and} \quad H = - \frac{h^2}{8\pi^2 m} \Delta,$$

so:

$$- \frac{h^2}{8\pi^2 m} \Delta \varphi = E \varphi.$$

Let \mathbf{p} be the impulse vector of the corpuscle. One finds the proper functions:

$$\varphi(x, y, z, \mathbf{p}) = a e^{-\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} = a e^{-\frac{2\pi i}{h}(\mathbf{p} \cdot \mathbf{r})},$$

with:

$$(I.c) \quad \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) = \frac{p^2}{2m} = E.$$

One then sees that:

1. Any positive value of E is a proper value.

2. Any proper value of E corresponds to an infinitude of proper functions of the preceding type that are obtained by giving p_x, p_y, p_z all possible values that are compatible with (I.c).

Hence, one will find a continuous spectrum for energy that goes from 0 to $+\infty$, with degeneracy of order ∞ for each value of E other than 0.

Each proper function corresponds to a monochromatic plane wave of the form:

$$\psi_{\mathbf{p}}(x, y, z, t) = \varphi(x, y, z, \mathbf{p}) e^{\frac{2\pi i}{h} E t} = a e^{\frac{2\pi i}{h}(E t - \mathbf{p} \cdot \mathbf{r})}.$$

One can simplify the writing by setting:

$$k = \frac{2\pi}{hc} E, \quad \mathbf{k} = \frac{2\pi}{hc} \mathbf{p},$$

which will give:

$$\psi_{\mathbf{k}} = a e^{i(kct - \mathbf{k} \cdot \mathbf{r})},$$

with the relation:

$$kc = \frac{1}{2m} |\mathbf{k}|^2 \frac{h}{2\pi},$$

which will express (I.c) in the new notations.

The vector \mathbf{k} is called the *propagation vector* of the plane wave, which is specified entirely by that one given. One must indeed distinguish \mathbf{k} from $|\mathbf{k}|^2$.

One can just as well take the proper functions of H to be either the $\psi_{\mathbf{k}}$ or the $\varphi_{\mathbf{k}}$, which differ from them by the factor e^{ikct} .

One can express the orthonormality of the plane waves by introducing proper differentiation. In the course of that calculation (whose details we shall not reproduce),

we will be led to introduce, with Dirac, the “improper” or “singular” function $\delta(x)$, which has the following properties:

1. $\delta(x)$ is an even function of x .
2. One always has $\int_{x_1}^{x_2} f(x) \delta(x) dx = \begin{cases} f(0) & \text{if } x_1 \text{ and } x_2 \text{ have opposite signs} \\ 0 & \text{if } x_1 \text{ and } x_2 \text{ have the same sign.} \end{cases}$

One can represent $\delta(x)$ in terms of Dirichlet’s singular function by setting:

$$\delta(x) = \lim_{N \rightarrow \infty} \frac{\sin 2\pi Nx}{\pi x}.$$

Finally, the calculation in question will show that the proper functions of the continuous spectrum must be written:

$$\varphi(\mathbf{k}, x, y, z) = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{r}}, \quad \psi(\mathbf{k}, x, y, z) = \frac{1}{(2\pi)^{3/2}} e^{i(kct - \mathbf{k}\cdot\mathbf{r})}.$$

The complete character of the set of $\varphi_{\mathbf{k}}$ translates into the fact that under very general conditions a function $f(x, y, z)$ can be developed into a Fourier integral in the form:

$$f(x, y, z) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} c(\mathbf{k}) e^{+i\mathbf{k}\cdot\mathbf{r}} d\mathbf{k},$$

in which $d\mathbf{k}$ denotes dk_x, dk_y, dk_z . The $c(\mathbf{k})$ are given by the formula:

$$c(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int_D f(x, y, z) e^{+i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r},$$

in which $d\mathbf{r}$ denotes dx, dy, dz . That is the inversion formula that gives the coefficients of the Fourier development.

One can also write:

$$f(x, y, z) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} c(\mathbf{k}, t) \psi_{\mathbf{k}}(x, y, z, t) d\mathbf{k},$$

with:

$$c(\mathbf{k}, t) = c(\mathbf{k}) e^{-ikct}.$$

CHAPTER II

PHYSICAL INTERPRETATION OF WAVE MECHANICS

1. GENERAL PRINCIPLES

We have seen that wave mechanics must be able to calculate the possible values of measurable quantities that are attached to a corpuscle and their respective probabilities. We have learned to represent the state of a corpuscle by a wave function $\psi(x, y, z, t)$ that is a solution to the equation of propagation, which is a function that we will always assume to have been normalized. Moreover, we have made any quantity that is attached to a corpuscle correspond to a linear, Hermitian operator that permits one to define a set of real numbers – viz., its proper values – and a complete system of orthonormal functions – viz., its proper functions. We are then in a position to state the two fundamental principles of the physical interpretation of wave mechanics.

First principle: The possible values of a measurable quantity – i.e., the various possible results of a measurement of that quantity – will be the proper values of the linear, Hermitian operator that corresponds to that quantity.

Second principle: When the state of a corpuscle is represented by a certain wave function $\psi(x, y, z, t)$ that is a solution of the equation of propagation, the probability that a precise measurement of the measurable quantity that corresponds to the linear, Hermitian operator A , which is complete and has non-degenerate proper values, will yield a certain proper value at the instant t is equal to the square of the modulus of the coefficient of the corresponding proper function in the development of the function ψ in proper functions of A .

More precisely, if the function ψ is developed into proper functions and proper differentials of A according to the formula:

$$\psi(x, y, z, t) = \sum_i c_i \varphi_i + \sum_{\Delta\alpha} c(\alpha) \int_{\alpha}^{\alpha+\Delta\alpha} \varphi(\alpha) d\alpha$$

then the probability of the proper value α_i is $|c_i|^2$, and the probability that a value will be found in the interval $(\alpha, \alpha + \Delta\alpha)$ is $|c(\alpha)|^2 \Delta\alpha$. One can verify that the function ψ is normalized, by hypothesis, so the total probability of all the possible hypotheses will be equal to unity. Naturally, the probability of a possible value can be a function of time.

If the operator A has multiple proper values then the statement of the second principle must be modified. Let α_i be a multiple proper value that corresponds to p proper functions $\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{ip}$ that are normalized, orthogonal, and linearly-independent. The probability of finding the value α_i for the quantity A from a measurement that is made at the instant t will be the sum of the squares of the moduli of

the coefficients of $\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{ip}$ in the development of ψ , namely, $\sum_{j=1}^p |c_{ij}|^2$. One proves that the value of that probability is independent of the fashion by which the φ_{ij} are chosen (which is, to a certain extent, arbitrary). In other words, the quantity $\sum_{j=1}^p |c_{ij}|^2$ will remain invariant when one replaces the φ_{ij} with p linearly-independent linear combinations of the φ_{ij} .

When the operator A is incomplete, the statement of the second principle must submit to another modification. Indeed, the proper functions of A will not contain all of the variables x, y, z then, and the coefficients c_i and $c(\alpha)$ will be functions that are not addressed by A . The probability of a proper value α_i cannot be the corresponding $|c_i|^2$, which will further depend upon certain variables. In order to obtain that probability, one must integrate over those variables. With that modification one can verify that the total probability of all the possible values is in fact equal to one.

As E. Arnous has shown, the two general principles of the physical interpretation of wave mechanics can be summarized in a single statement. In order to do that, it will suffice to assume the following postulate: The probability distribution that corresponds to the measurable values of the observable quantity A has the characteristic function:

$$\varphi(u) = \int_D \psi^* e^{iAu} \psi d\tau,$$

in which A is the operator that corresponds to the quantity A . We shall not insist upon that very interesting and elegant formulation of the laws of probability in wave mechanics here.

A simple example of an application of the two principles is provided by the case of the Hamiltonian operator H , which is complete. If H is independent of time then it will take on constant proper values E_i and proper functions φ_i . A precise measurement of the energy can provide only one of the values E_i , and if one has $\psi = \sum_i c_i \varphi_i$ then the probability of a value E_i will be $|c_i|^2$. If the spectrum is discrete then one will have a discrete sequence of stationary states with quantized energies. That is the case that presents itself for the atomic systems.

Take another case: viz., that of the coordinate x of the corpuscle that corresponds to the operation of “multiplication by x .” The proper value equation is:

$$\kappa \varphi = \alpha \varphi.$$

That equation can be considered to be verified for any real value of x by setting:

$$\varphi(x, \alpha) = \delta(x - \alpha),$$

in which $\delta(x - \alpha)$ is Dirac’s singular function for $x - \alpha$. Hence, from the first principle, a measurement of x can yield any value of x that is real and found between $-\infty$ and $+\infty$.

Moreover, the proper differentials of that continuous spectrum $\int_{\alpha}^{\alpha+\Delta\alpha} \delta(x-\alpha) d\alpha$ will define a complete system that satisfies the orthonormality relation. Since one obviously has:

$$\psi(x, y, z, t) = \int_{-\infty}^{+\infty} \psi(\alpha, y, z, t) \delta(x-\alpha) d\alpha,$$

the probability that a measurement of x at the instant t will yield a value that is found in the interval $(\alpha, \alpha + \Delta\alpha)$ will be:

$$\int_{-\infty}^{+\infty} dy dz |\psi(\alpha, y, z, t)|^2 \Delta\alpha.$$

One easily deduces that the probability of presence for the corpuscle to manifest itself around the point x, y, z in the volume element $d\tau$ at the instant t will be $|\psi(x, y, z, t)|^2 d\tau$. The total probability of presence at an arbitrary point in space is indeed equal to unity, since ψ is normalized. That is the physical reason why the wave function must always be normalized. The interpretation that was given to $|\psi|^2$ is then in accord with the positive-definite character of that quantity ($|\psi|^2 \geq 0$).

One infers the following conclusion from the two fundamental principles by an argument upon which we shall not insist: Two measurable quantities can be measured simultaneously with precision by the same measurement operation only if the corresponding operators A and B commute; i.e., only if $AB = BA$.

The most important example of two quantities that are simultaneously measurable is that of a coordinate x and the conjugate component of the quantity of motion p_x . Indeed, one will have:

$$(x p_x - p_x x)_{\text{op}} = \frac{h}{2\pi i} \left(\frac{\partial}{\partial x} x - x \frac{\partial}{\partial x} \right) = \frac{h}{2\pi i} \neq 0.$$

Hence, a Lagrange coordinate and its conjugate momentum are not simultaneously measurable with precision. Their simultaneous measurements are affected with “uncertainties” that cannot be simultaneously zero. Later on, we will see that one always has $\Delta x \cdot \Delta p_x \geq h$ for those uncertainties in order of magnitude.

Those are *Heisenberg's uncertainty inequalities*. More generally, for two canonically-conjugate quantities p and q , one will have:

$$\Delta p \cdot \Delta q \geq h.$$

With Bohr, one says that the quantities p and q correspond to “complementary” principles of reality that one cannot always know exactly at the same time.

One can say that when two quantities p and q are canonically-conjugate, one must make p correspond to the operator $-\frac{h}{2\pi i} \frac{\partial}{\partial q}$, and conversely. As an example, recall that

the kinetic moment of a corpuscle around an axis Oz is canonically-conjugate to the azimuth angle φ around that axis. Now, we have already shown that in wave mechanics,

one must make the classical quantity M_z correspond to the operator $-\frac{h}{2\pi i} \frac{\partial}{\partial \varphi}$, which confirms the preceding statements.

2. ALGEBRAIC MATRICES AND THEIR PROPERTIES

One says *matrix* to mean an array of numbers that contains a finite or infinite number of rows and columns. If that array has finite dimensions then we shall assume that it is square, for the sake of simplicity. Let a_{ik} be the element of the matrix that is found at the intersection of the i^{th} row with the j^{th} column. The elements a_{ii} with equal indices are called the *diagonal* elements; a matrix whose only non-zero elements are diagonal is called a *diagonal matrix*. Two matrices A and B are said to be equal if $a_{ik} = b_{ik}$ for all i and k .

Matrices present themselves in algebra when one studies linear transformations. Indeed, if the variables x'_i are linear combinations of other variables x_i then one will have transformation formulas of the type:

$$x'_i = \sum_j a_{ij} x_j,$$

or symbolically:

$$x' = Ax,$$

with the convention that:

$$(AX)_i = \sum_j a_{ij} x_j.$$

Those formulas lead one to define the sum and product of two matrices by the following rules:

1. The sum of A and B is the matrix $A + B$, whose element with the indices i, k will be $a_{ik} + b_{ik}$.

2. The product of A by B is the matrix AB whose element ik is $\sum_j a_{ij} b_{jk}$.

It results from that definition that $AB \neq BA$, in general. If $AB = BA$, exceptionally, then one says that the matrices *commute*. One refers to the matrix $[A, B] = AB - BA$ as the *commutator* of A and B , which measures the lack of commutation of A and B when it is not zero. Sometimes one also introduces the *anti-commutator* of A and B that is defined by $[A, B]_+ = AB + BA$. If $AB = -BA$ then A and B *anti-commute*.

Matrices can be real or complex, according to whether their elements are real or complex, resp. We shall envision the general case of complex matrices.

A matrix is called *Hermitian* if one has:

$$a_{ik} = a_{ki}^*$$

for any i and k . A real Hermitian matrix is then symmetric with respect to its diagonal. The diagonal elements of a Hermitian matrix are real.

A matrix is *anti-Hermitian*, and one will have:

$$a_{ik} = -a_{ki}^*$$

for any i and any k . The diagonal elements of an anti-Hermitian matrix are pure imaginary.

The product of two Hermitian matrices A and B is itself Hermitian if A and B commute, and only in that case; it is anti-Hermitian only if A and B anti-commute.

The matrix \tilde{A} is the *transposed* matrix to A if $\tilde{a}_{ik} = a_{ki}$. One says that A^\dagger is the *adjoint* matrix to A if $a_{ik}^\dagger = a_{ki}^*$; hence, $A^\dagger = \tilde{A}^*$. If A is Hermitian then one will have $A = A^\dagger$; the matrix A will then be its own adjoint. One obviously has $(A^\dagger)^\dagger = A$, and one easily proves that $(AB)^\dagger = B^\dagger A^\dagger$.

A Hermitian matrix that is diagonal is necessarily real. The matrix whose elements are $a_{ik} = \delta_{ik}$ is diagonal, and all of its diagonal terms are equal to 1. If one is given a matrix A and there exists a matrix A^{-1} such that:

$$AA^{-1} = A^{-1}A = 1$$

then the matrix A^{-1} will be called *inverse* to A . If A has a finite number of rows and columns then A^{-1} will always exist when the determinant that is deduced from a_{ik} is non-zero. If A has an infinite number of rows and columns then A^{-1} might or might not exist, according to the case. One easily proves that $(AB)^{-1} = B^{-1}A^{-1}$. When A is a real matrix and one has:

$$\sum_i a_{ij} a_{ik} = \delta_{jk}, \quad \sum_i a_{ji} a_{ki} = \delta_{jk},$$

one says that the matrix is *orthogonal*. It will then define an *orthogonal* transformation, which leaves the sum $\sum_i x_i^2$ invariant, as is well-known in geometry. One generalizes that definition for a complex matrix A by saying that if one has:

$$\sum_i a_{ij} a_{ik}^* = \delta_{jk}, \quad \sum_i a_{ji} a_{ki}^* = \delta_{jk}$$

then the matrix will define a complex orthogonal transformation [*sic*] or *unitary* transformation. For such a transformation, the sum of the norms $\sum_i x_i x_i^*$ of the x_i will remain invariant, as one easily proves. For a *unitary* matrix A , one will have:

$$\sum_i a_{ki}^* a_{ij} = \delta_{kj}, \quad \sum_i a_{ji} a_{ik}^* = \delta_{jk},$$

or furthermore:

$$A^\dagger A = AA^\dagger = 1, \quad \text{so} \quad A^\dagger = A^{-1}.$$

Hence, the adjoint of a unitary matrix will coincide with its inverse.

The *trace* of a matrix A is the sum of its diagonal terms:

$$\text{Tr}(A) = \sum_i a_{ii}.$$

One easily proves that:

$$\text{Tr}(AB) = \text{Tr}(BA) = \sum_{i,k} a_{ik} b_{ki}.$$

Now, let A be an arbitrary square matrix, and let S be a unitary matrix that has the same number of rows and columns. The matrix $S^{-1}AS = B$ is said to be obtained from A by a *canonical transformation*. One easily verifies that if A is Hermitian then B will also be so. Canonical transformations will then preserve the Hermitian character of a matrix. It is easy to verify that they will also preserve the trace. Moreover, if two square matrices A and A' transform into B and B' , respectively, then their product will transform into BB' under that transformation, because $S^{-1}ASS^{-1}A'S = S^{-1}AA'S$.

3. OPERATORS AND MATRICES IN WAVE MECHANICS

Suppose that we know a system of orthonormal functions $\varphi_1, \dots, \varphi_i, \dots$ in a domain D in which certain variables vary; we call them *basis functions*. That system can be the system of normalized proper functions of a linear, Hermitian operator in wave mechanics. With that basis system, any linear operator will correspond to a matrix. Indeed, let A be a linear operator. The application of that operator to one of the basis functions φ_i will yield a new function, which can be developed in the same system of basis functions. We will then have a relation of the form:

$$A\varphi_i = \sum_j a_{ji} \varphi_j,$$

with:

$$a_{ji} = \int_D \varphi_j^* A\varphi_i d\tau,$$

in which D is the domain of variation of the variables that enter into φ_i . By definition, the a_{ij} are elements of the matrix that is generated by the operator A in the basis system. We denote that by the same symbol A that we use for the operator, or by A^φ , if we wish to specify the basis system that is employed. It is easy to verify that the matrices thus-defined will verify the rules of addition and multiplication of algebraic matrices that were studied above.

If the basis system is defined by the proper functions of an operator in wave mechanics, and if A is itself a linear, Hermitian operator in that mechanics then we will say that A is a *matrix of wave mechanics*. One will see immediately that those matrices are always themselves Hermitian; i.e., that $a_{ij} = a_{ji}^*$.

One sees, moreover, that the necessary and sufficient condition for the matrix that is generated by an operator A in a basis system to be Hermitian is that the operator is itself Hermitian. Hermiticity is then an intrinsic property of operators, in the sense that a

Hermitian operator will generate Hermitian matrices in all systems of basis functions. All of the matrices of wave mechanics will be Hermitian.

Our definitions establish a very close correlation between operators and matrices. In particular, the necessary and sufficient condition for the matrices to commute (or anti-commute) is that the corresponding operators must commute (or anti-commute), and *vice versa*. That leads us to define the commutator or anti-commutator of two operators A and B by formulas:

$$[A, B] = AB - BA, \quad [A, B]_+ = AB + BA.$$

A very important category of matrices in wave mechanics is obtained by always taking the basis functions to be the proper functions of the Hamiltonian operator that corresponds to the problem considered. Let $\psi_1, \dots, \psi_i, \dots$ be the proper functions of the operator H . The matrices A that are generated by a linear, Hermitian operator in the system of ψ_i , whose elements will be:

$$a_{jk} = \int_D \psi_j^* A \psi_k d\tau,$$

can be called the matrices of quantum mechanics, because they are the ones that Heisenberg put at the basis of quantum mechanics without interpreting them explicitly. If one includes the exponential factors $e^{\frac{2\pi i}{h} E_k t}$ in the ψ – i.e., if one sets:

$$\psi_k = a_k(x, y, z) e^{\frac{2\pi i}{h} E_k t}$$

then one will have:

$$a_{jk} = \int_D \psi_j^* A \psi_k d\tau \cdot e^{\frac{2\pi i}{h} (E_k - E_j) t}.$$

Those elements define the Heisenberg matrices, properly speaking, which depend upon time. Sometimes, one suppresses the exponential factor in the expression for ψ_k and sets simply:

$$a'_{jk} = \int_D \psi_j^* A \psi_k d\tau.$$

The matrix A' of elements a'_{jk} that are independent of time is called the *Schrödinger matrix* that corresponds to the operator A .

4. MEAN VALUES AND FIELD QUANTITIES IN WAVE MECHANICS

Imagine a certain corpuscle and suppose that the wave ψ that it is associated with is known. Let a certain observable quantity be attached to that corpuscle that corresponds to a linear, Hermitian operator A . The principles of wave mechanics will then permit us to predict the possible observed values for A and their probabilities. Since there are, in general, several possible values with non-zero probabilities, one cannot speak of the

value of A at each instant, but one can easily define its mean value (i.e., mathematical expectation) by the formula:

$$\bar{A} = \sum_i |c_i|^2 \alpha_i .$$

It is easy to verify that one has, equivalently:

$$\bar{A} = \int_D \psi^* A \psi d\tau .$$

Now let B be an observable quantity other than A that is attached to the corpuscle and corresponds to a linear, Hermitian operator B . Let $\beta_1, \dots, \beta_k, \dots$ and $\chi_1, \dots, \chi_k, \dots$ be the proper values and proper functions of B . If the function ψ is developed in the χ_i by the formula $\psi = \sum_i d_i \chi_i$ then one will have:

$$\bar{A} = \sum_{i,k} d_i^* d_k a_{ik}^* ,$$

in which a_{ik}^* is the element with indices i, k of the matrix that is generated by A in the system of χ .

Hence, the mean value of A can always be expressed linearly with the aid of the matrix elements that the operator A generates in the system of proper function of another operator B .

In particular, if the corpuscle is found in one of the proper states that relate to the quantity B then one will have $\psi = d_i \chi_i$, with $|d_i| = 1$, and in turn $\bar{A} = a_{ii}^*$. One will then have the theorem:

The diagonal element with indices i, i in the matrix that is generated by the operator A in the system of proper functions of the operator B is equal to the mean value of the quantity A when one knows that B has the value χ_i .

The system of Heisenberg matrices has the peculiarity that the matrix H that corresponds to the energy is represented by a diagonal matrix whose diagonal terms are the various proper values of energy. (In the case where H has multiple proper values, one must take care to choose the corresponding proper functions in such a fashion that they will be orthogonal.) Indeed, one has:

$$H_{ik} = \int_D \psi_i^* H \psi_k d\tau = E_k \int_D \psi_i^* \psi_k d\tau = E_k \delta_{ik} .$$

That result is a particular case of the general result that:

If one constructs the matrix that is generated by an operator A in the system of orthonormal proper functions of that operator then that matrix will be diagonal, and its diagonal elements will be the proper values of A .

In the definitions of the matrix elements $a_{ik} = \int_D \psi_i^* A \psi_j d\tau$ and the mean values $\bar{A} = \int_D \psi^* A \psi d\tau$, the quantities under the sign \int_D are functions of x, y, z , and possibly the parameter t . We call them *densities of the matrix elements* or *mean-value densities*, resp. Those densities vary from one point to another in space and have the character of field quantities that are attached to the corpuscle. Hence, the quantity A can be associated with a field quantity, namely, the mean density of A :

$$\rho(A) = \psi^* A \psi.$$

Nevertheless, the field quantities thus-defined will not have a physical sense that is as precise as in the classical theories of physical fields. Here, they present themselves as being only “quantities that one must integrate in order to obtain the mean values (or the matrix elements).” Those are often complex quantities, and they are defined only up to a divergence, moreover. Meanwhile, they are quantities that are physically ill-defined from the quantum viewpoint, which will be quantities with a well-defined relativistic variance in quantum theories like that of Dirac.

5. FIRST INTEGRALS IN WAVE MECHANICS

The matrix element $a_{jk} = \int_D \psi_j^* A \psi_k d\tau$ can depend upon the parameter t by the intermediary of ψ_j, ψ_k , and also A , if that operator contains t in its definition. The derivative of a_{jk} , which is calculated by taking into account the facts that ψ_j and ψ_k are solutions of the wave equation and A is Hermitian, is:

$$\frac{da_{jk}}{dt} = \int_D \psi_j^* \left[\frac{\partial A}{\partial t} + \frac{2\pi i}{h} (AH - HA) \right] \psi_k d\tau,$$

in which $\partial A / \partial t$ is the operator that is obtained by formally differentiating A with respect to t . One can say that the Heisenberg matrix whose element with indices j, k is da_{jk} / dt is generated by the operator $\frac{\partial A}{\partial t} + \frac{2\pi i}{h} (AH - HA)$, and one sets:

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \frac{2\pi i}{h} (AH - HA) = \frac{\partial A}{\partial t} + \frac{2\pi i}{h} [A, H].$$

If A does not depend upon t explicitly (which is a frequent case) then one will have simply:

$$\frac{dA}{dt} = \frac{2\pi i}{h} [A, H].$$

By definition, in a problem where the Hamiltonian H is given, the observable quantity A is a *first integral* or *constant of the motion* for the problem in question if $dA / dt = 0$. If A does not depend upon time explicitly then A will be a first integral if A and H commute.

One can also define first integrals in the following fashion: A quantity whose operator is A will be a first integral if the fact that ψ is an arbitrary solution of the wave equation implies that $A\psi$ is likewise one. Indeed, if $\frac{\partial \psi}{\partial t} = \frac{2\pi i}{h} AH\psi$, by hypothesis, then one will have $A \frac{\partial \psi}{\partial t} = \frac{2\pi i}{h} AH\psi$ and $\frac{\partial}{\partial t} A\psi = \frac{\partial A}{\partial t} \psi + \frac{2\pi i}{h} AH\psi$. In order for $A\psi$ to be a solution of the wave equation, it is necessary that one must have:

$$\frac{\partial A}{\partial t} \psi + \frac{2\pi i}{h} (AH - HA) \psi = 0.$$

The necessary and sufficient condition for that to be true for any solution ψ of the wave equations is precisely that $dA / dt = 0$. Q. E. D.

Here are some examples of first integrals: If the external field that acts upon a corpuscle is independent of time then the operator H will not contain t , and since it obviously commutes with itself, energy will be a first integral. We then recover the analogue of the conservation of energy for the conservative systems of classical mechanics. Similarly, if the x -component of the field is zero then the operator H will not depend upon x , and in turn, it will commute with $(p_x)_{op}$. The x -component of the quantity of motion will then be a first integral, which is a theorem that is analogous to a theorem in classical mechanics.

Finally, if the function U possesses cylindrical symmetry around Oz , then the Hamiltonian H will not depend upon the azimuth ϕ around that axis. In that case, the component M_z of the kinetic moment around Oz is a first integral. If the force is central then the three components M_x, M_y, M_z will be first integrals, and the same thing will be true for the quantity $M^2 = M_x^2 + M_y^2 + M_z^2$. We shall return to this important case much later.

6. PRECISE FORM FOR THE UNCERTAINTY RELATIONS

We shall give a precise form to the Heisenberg uncertainty relations that is due to Pauli. One should note that this precise form is not completely equivalent to the qualitative form $\delta q \cdot \delta p \geq h$, which often has a physical significance that is more directly accessible to the experimenter, and which can even be valid when the precise form is not.

We begin by introducing the following definition: We say that the operator F^\dagger is the *adjoint operator* to F in a domain D if:

$$\int_D f^* F g d\tau = \int_D (F^\dagger f)^* g d\tau,$$

in which f and g are functions on the domain D that are subject to only the conditions that they must be finite, uniform, and continuous on D , and that they should be annulled on the boundary of D in such a fashion that the surface integrals that might appear in the integration by part of \int_D will be zero. Upon comparing the definition of the adjoint operator with that of the Hermitian operators, one will see that an operator is Hermitian if and only if it is its own adjoint (i.e., $F = F^\dagger$).

Whether the operator F is Hermitian or not, the mean value of FF^\dagger will always be real and positive-definite because:

$$(II.a) \quad \overline{FF^\dagger} = \int_D \psi^* FF^\dagger \psi d\tau = \int_D (F^\dagger \psi)^* F^\dagger \psi d\tau = \int_D |F^\dagger \psi|^2 d\tau.$$

Having said that, we shall prove the following theorem:

Theorem: *If two observable physical quantities correspond to the linear, Hermitian operators A and B , respectively, then one will have:*

$$\sigma_A \cdot \sigma_B \geq \frac{1}{2} \overline{[A, B]},$$

in which $[A, B]$ is the commutator of A and B , and σ_A , σ_B are the standard deviations (i.e., dispersions) that are defined by:

$$\sigma_A = \sqrt{\overline{(A - \bar{A})^2}}, \quad \sigma_B = \sqrt{\overline{(B - \bar{B})^2}}.$$

In order to prove that theorem, we consider the linear, non-Hermitian operator $A + i\lambda B$, where λ is a real constant. Its adjoint is $A - i\lambda B$, and upon applying formula (II.a), we will see that:

$$\overline{(A + i\lambda B)(A - i\lambda B)} = \overline{A^2 + \lambda^2 B^2 - i\lambda [A, B]}$$

is real and positive-definite. Hence, the function of λ :

$$f(\lambda) = \overline{A^2 + \lambda^2 B^2 - i\lambda [A, B]}$$

is real and positive-definite. One then concludes that $\overline{[A, B]}$ is pure imaginary. Now,

$f(\lambda)$ has a minimum for $\lambda_0 = \frac{1}{2} \frac{\overline{[A, B]}^2}{\overline{B^2}}$, and one will then have the value $f(\lambda_0) = \overline{A^2} + \frac{1}{4} \frac{\overline{[A, B]}^2}{\overline{B^2}}$. Since that value must be positive or zero, one will have:

$$(II.b) \quad \overline{A^2} \cdot \overline{B^2} \geq -\frac{1}{4} \overline{[A, B]}^2.$$

Set $\Delta A = A - \bar{A}$ and $\Delta B = B - \bar{B}$, by definition. \bar{A} and \bar{B} are numbers, but since A and B are operators, ΔA and ΔB will also be operators, and one will easily find that:

$$(II.c) \quad [\Delta A, \Delta B] \equiv [A - \bar{A}, B - \bar{B}] = [A, B].$$

Upon applying the inequality (II.b) to the operators ΔA and ΔB and taking (II.c) into account, that will give:

$$\overline{(\Delta A)^2} \cdot \overline{(\Delta B)^2} \geq -\frac{1}{4}(\overline{[A, B]})^2.$$

Since $\overline{[A, B]}$ is pure imaginary, one can infer that:

$$\sigma_A \cdot \sigma_A = \sqrt{\overline{(\Delta A)^2}} \cdot \sqrt{\overline{(\Delta B)^2}} \geq \frac{1}{2} \overline{[A, B]},$$

and the theorem is proved.

In wave mechanics, one says two observable quantities A and B are *canonically conjugate* when:

$$[A, B] = -\frac{h}{2\pi i} 1.$$

One will then have:

$$\overline{[A, B]} = \int_D \psi^* \left(-\frac{h}{2\pi i} \right) \psi d\tau = -\frac{h}{2\pi i},$$

which is a pure imaginary quantity, as it must be. That will give us:

$$\sigma_A \cdot \sigma_A \geq \frac{h}{4\pi}.$$

That is the precise form of the uncertainty relations that was promised, and it gives:

$$\sigma_x \cdot \sigma_{p_x} \geq \frac{h}{4\pi},$$

in particular.

CHAPTER III

**QUANTUM THEORY
OF KINETIC MOMENTS AND SPINS**

1. ORBITAL KINETIC MOMENT

We say the “orbital kinetic moment” (i.e., moment of impulse or moment of rotation) of a particle with respect to a point O , which is taken to be the coordinate origin, to mean the moment of the quantity of motion of the particle with respect to that point. That orbital kinetic moment is a vector whose expression is:

$$\mathbf{M} = [\mathbf{r} \times \mathbf{p}].$$

It is then the vector product of the radius vector \mathbf{r} of the particle (with components x, y, z) and the quantity of motion \mathbf{p} . In components, one will then have:

$$M_x = y p_z - z p_y, \quad M_y = z p_x - x p_z, \quad M_z = x p_y - y p_x.$$

The essential property of the kinetic moment that will make that quantity particularly important from the standpoint of mechanics is that if the potential of the forces that act upon the particle does not depend upon the azimuth that is taken around one of the coordinate axes (in other words, if the force is everywhere in the same plane as that axis) then the component of the orbital kinetic moment along that axis will be constant in the course of motion; in other words, it is a first integral.

The length M of the orbital kinetic moment is defined by:

$$M^2 = M_x^2 + M_y^2 + M_z^2 = r^2 p^2 - (\mathbf{r} \cdot \mathbf{p})^2,$$

from the Lagrange identity.

\mathbf{M} will be a first integral if the force constantly passes through the point O (i.e., a central force).

In wave mechanics, we must replace the quantities that were defined that way classically with operators. We must set:

$$(M_x)_{\text{op}} = -\frac{h}{2\pi i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) = -\frac{h}{2\pi i} \frac{\partial}{\partial \varphi_x}, \quad (M_y)_{\text{op}} = -\frac{h}{2\pi i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) = -\frac{h}{2\pi i} \frac{\partial}{\partial \varphi_y},$$

$$(M_z)_{\text{op}} = -\frac{h}{2\pi i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) = -\frac{h}{2\pi i} \frac{\partial}{\partial \varphi_z}.$$

In those expressions, $\varphi_x, \varphi_y, \varphi_z$ are the azimuths when measured around the axes Ox, Oy, Oz , resp. Any of those three operators $M_k = \frac{-h}{2\pi i} \frac{\partial}{\partial \varphi_k}$, with $k = 1, 2, 3$, will admit proper values that are equal to $hm / 2\pi$, in which m is an integer (positive, negative or zero), and normalized proper functions that are equal to $\frac{1}{\sqrt{2\pi}} e^{-im\varphi_k}$, as one easily verifies.

From the general principles of wave mechanics, one must conclude that the exact measurement of one of the rectangular components of the kinetic moment will always give a value that is equal to an integer multiple of $h / 2\pi$. For that reason, $h / 2\pi$ can be called the *quantum unit* of kinetic moment.

One then perceives that the vectorial image of the kinetic moment that is provided by the classical theory is somewhat misleading at the quantum level. Indeed, the three components of the kinetic moment are not simultaneously measurable at the quantum level, because the operators M_x, M_y, M_z do not commute with each other. If one then performs a precise measurement of one of those components then one will get only a probability distribution for the values of the other two, so one cannot know the exact value.

One cannot actually trace out the vector \mathbf{M} , since never knows more than one component exactly. By contrast, the vector $\bar{\mathbf{M}}$, whose components are the mean values $\bar{M}_x, \bar{M}_y, \bar{M}_z$, is always well-defined, and that is what permits one to employ an orbital moment vector at the macroscopic level, at which one can measure only its mean value.

The impossibility of knowing the three components of the kinetic moment simultaneously is expressed by the non-commutation of the corresponding operators. In order to describe the non-commutation formulas, we explicitly include the unit $h / 2\pi$.

$$M_x = m_x \cdot \frac{h}{2\pi}, \quad M_y = m_y \cdot \frac{h}{2\pi}, \quad M_z = m_z \cdot \frac{h}{2\pi},$$

with:

$$m_x = i \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \quad m_y = i \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \quad m_z = i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

One can easily verify that:

$$(III.a) \quad [m_x, m_y] = -i m_z, \quad [m_y, m_z] = -i m_x, \quad [m_z, m_x] = -i m_y.$$

We shall see the significance of those formulas later on.

We also set:

$$(M^2)_{\text{op}} = (M_x)_{\text{op}}^2 + (M_y)_{\text{op}}^2 + (M_z)_{\text{op}}^2 = \frac{h^2}{4\pi^2} (m_x^2 + m_y^2 + m_z^2) = m^2 \frac{h^2}{4\pi^2},$$

and we will find that:

$$m^2 = \frac{1}{\sin \theta} \left[\frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right],$$

when one takes polar coordinates around O . The operator m^2 is nothing but the Laplacian on the surface of a sphere of radius 1. The proper value equation $m^2 f = \alpha f$ admits only the Laplace functions $y_k(\theta, \varphi)$ for solutions that are finite, continuous, and uniform on the sphere of radius 1, and the proper value that corresponds to the proper function y_l is $l(l + 1)$ where l is an integer that is positive or zero. Finally, the proper values of the operator M^2 are:

$$M_l^2 = l(l + 1) \frac{h^2}{4\pi^2} \quad (l = 0, 1, 2, \dots).$$

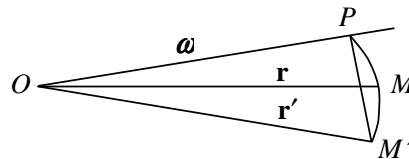
It is easy to verify that $(m^2)_{\text{op}}$ commutes with m_x, m_y, m_x , which shows that M^2 is a measurable quantity at the same time as the quantities M_x, M_y, M_z .

2. KINETIC MOMENT AND THE GROUP OF SPATIAL ROTATIONS

In order to better comprehend the sense of the non-commutativity of the components of the kinetic moment, it is useful to show how the character of M_z that it is a first integral relates to rotations around Oz . In order for M_z to be a first integral, it is necessary and sufficient that the potential U does not depend upon the azimuth φ around Oz . But then, the problem in wave mechanics will not be altered in any way by a rotation of the system around Oz through an arbitrary angle $\Delta\varphi$. Hence, if $\psi(r, \theta, \varphi, t)$ is the expression for a solution of the wave equation in polar coordinates with Oz for the polar axis then $\psi(r, \theta, \varphi + \Delta\varphi, t)$ will also be a solution, and the same thing will be true for $\frac{\psi(r, \theta, \varphi + \Delta\varphi, t) - \psi(r, \theta, \varphi, t)}{\Delta\varphi}$, which will be equal to $\partial\psi / \partial\varphi$. Hence, the hypothesis

that the problem is not modified by a rotation around Oz will imply that if ψ is a solution of the wave equation then $\partial\psi / \partial\varphi$, and in turn, $M_z\psi$, will also be one. We have seen that this result implies that M_z will have the character of a first integral. Hence, we easily shed light upon the relationship between M_z and the rotations around Oz .

We shall now show that the non-commutation of the components of the kinetic moment is related to the non-commutation of spatial rotations. Pick a point M , whose coordinates x, y, z form the components of a radius vector $\overline{OM} = \mathbf{r}$, and let an infinitesimal rotation be defined by a vector $\boldsymbol{\omega}$ that passes through O . Under the influence of the rotation $\boldsymbol{\omega}$ the point M will go to M' , while describing an infinitely-small arc MM' of a circle that is equal to $\omega \overline{MP}$ or $\omega r \sin \widehat{POM}$:



One will then have:

$$\mathbf{r}' = \mathbf{r} + [\boldsymbol{\omega} \times \mathbf{r}].$$

Now consider three infinitesimal rotations with the same absolute value (equal to 1) around the three rectangular axes Ox , Oy , Oz . They correspond to three vectors $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, $\boldsymbol{\omega}_3$ that are likewise carried by Ox , Oy , Oz , resp. One obviously has:

$$\boldsymbol{\omega}_1 = [\boldsymbol{\omega}_2 \times \boldsymbol{\omega}_3], \quad \boldsymbol{\omega}_2 = [\boldsymbol{\omega}_3 \times \boldsymbol{\omega}_1], \quad \boldsymbol{\omega}_3 = [\boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2].$$

Let $(\boldsymbol{\omega}_1 \boldsymbol{\omega}_2)_{\text{op}}$ denote the operation that consists of first applying the rotation $\boldsymbol{\omega}_2$, and then the rotation $\boldsymbol{\omega}_1$ to a point P , and let $(\boldsymbol{\omega}_2 \boldsymbol{\omega}_1)_{\text{op}}$ denote the operation that consists of applying those rotations to P , but in the opposite order. The operation $(\boldsymbol{\omega}_1 \boldsymbol{\omega}_2)_{\text{op}}$ takes the initial point P to a point P_1 such that:

$$\mathbf{r}_1 = \mathbf{r} + [\boldsymbol{\omega}_1 \times \mathbf{r}] + [\boldsymbol{\omega}_2 \times \mathbf{r}] + [\boldsymbol{\omega}_1 \times (\boldsymbol{\omega}_2 \times \mathbf{r})],$$

and the operation will take the point P to the point P_2 such that:

$$\mathbf{r}_2 = \mathbf{r} + [\boldsymbol{\omega}_2 \times \mathbf{r}] + [\boldsymbol{\omega}_1 \times \mathbf{r}] + [\boldsymbol{\omega}_2 \times (\boldsymbol{\omega}_1 \times \mathbf{r})].$$

The two operations are not equivalent, and will not give the same result; i.e., the rotations $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ do not commute. The difference between the results corresponds to the symbol $(\boldsymbol{\omega}_1 \boldsymbol{\omega}_2 - \boldsymbol{\omega}_2 \boldsymbol{\omega}_1)_{\text{op}}$, and is given by:

$$\mathbf{r}_1 - \mathbf{r}_2 = [\boldsymbol{\omega}_1 \times (\boldsymbol{\omega}_2 \times \mathbf{r})] - [\boldsymbol{\omega}_2 \times (\boldsymbol{\omega}_1 \times \mathbf{r})] = [\boldsymbol{\omega}_1 \times (\boldsymbol{\omega}_2 \times \mathbf{r})] + [\boldsymbol{\omega}_2 \times (\mathbf{r} \times \boldsymbol{\omega}_1)].$$

In a general fashion, one will have the easily-verified relation between three vectors \mathbf{A} , \mathbf{B} , \mathbf{C} :

$$[\mathbf{A} \times (\mathbf{B} \times \mathbf{C})] + [\mathbf{B} \times (\mathbf{C} \times \mathbf{A})] + [\mathbf{C} \times (\mathbf{A} \times \mathbf{B})] = 0.$$

Upon applying that relation to $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, and \mathbf{r} , one will get:

$$\mathbf{r}_1 - \mathbf{r}_2 = -[\mathbf{r} \times (\boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2)] = [(\boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2) \times \mathbf{r}] = [\boldsymbol{\omega}_3 \times \mathbf{r}].$$

The operation $(\boldsymbol{\omega}_1 \boldsymbol{\omega}_2 - \boldsymbol{\omega}_2 \boldsymbol{\omega}_1)_{\text{op}}$ is then equivalent to $\boldsymbol{\omega}_3$, which can be written in the form of an operator equation:

$$[\boldsymbol{\omega}_1 \boldsymbol{\omega}_2] = \boldsymbol{\omega}_3.$$

That relation is, in fact, equivalent to the one that corresponds to the first relation (III.a), because \mathbf{m}_k is equal to $i \partial / \partial \varphi_k$, and since a rotation with a certain sense will correspond to a variation of the values of ψ with the opposite sense, the symbol $i \partial / \partial \varphi_k$ will correspond to the symbol $-i (\boldsymbol{\omega}_k)_{\text{op}}$.

One finds analogous correspondences by circular permutation of x , y , z . Hence, the formulas for the non-commutation of the \mathbf{m}_k are found to be attached to the non-commutation of the rotation operators in space.

3. GENERAL RESULTS RELATING TO THE PROPER VALUES OF OPERATORS THAT SATISFY THE NON-COMMUTATION RELATIONS (III.a)

We shall now make a general study of the proper values of the three linear, Hermitian operators, which we assume to be subject to only the conditions that one must have:

$$[m_x, m_y] = -i m_z, \quad [m_y, m_z] = -i m_x, \quad [m_z, m_x] = -i m_y,$$

without supposing that they are equal to:

$$i \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \quad i \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \quad i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

We shall study the relationship between these proper values and those of the operator $m^2 = m_x^2 + m_y^2 + m_z^2$.

The proper values of all those operators are real, since the operators are Hermitian. The proper values of m_x^2 , m_y^2 , m_z^2 are therefore positive or zero, since they are the squares of the proper values of m_x , m_y , m_z . The mean values of m_x^2 , m_y^2 , m_z^2 are then necessarily positive, and the same thing will be true for those of m^2 . One concludes from this that the proper values of m^2 are positive or zero, since otherwise a proper state of m^2 that corresponds to a negative proper value of the mean value of m^2 would be negative, which is impossible.

By reason of symmetry, the proper values of m_x , m_y , m_z are the same, and we know that m^2 will commute with m_x , m_y , m_z . Since \mathbf{m}_x and \mathbf{m}_y are Hermitian, in each case, the mean value of m_z^2 cannot be greater than m^2 . Hence, the proper values of m_z^2 are less than or at most equal to the corresponding values of m^2 . In other words, the proper values of m_z cannot be either greater than m or less than $-m$.

Having said that, one easily verifies the relation:

$$(m_x - i m_y) m_z - m_z (m_x - i m_y) = - (m_x - i m_y),$$

so:

$$(m_x - i m_y) m_z = (m_z - 1) (m_x - i m_y).$$

Let γ_i and φ_i be the proper values and proper functions of m_z , resp. We shall appeal to the φ_i as a basis system for the construction of the matrices that we shall make use of. An operator A will then correspond to some matrix elements A_{ik}^φ , which we write simply as A_{ik} .

One will then have the following relation between the m_{ik} thus-constructed:

$$\sum_k (m_x - i m_y)_{ik} (m_z)_{kj} = \sum_k (m_z - 1)_{ik} (m_x - m_y)_{kj},$$

which is a translation of the last commutation relation, which will become simply:

$$(m_x - i m_y)_{ij} \gamma_j = (\gamma_i - 1)(m_x - i m_y)_{ij}$$

for the chosen basis system. For that relation to be verified, one must have:

$$\text{either: } \gamma_j = \gamma_j - 1 \quad \text{or: } (m_x - i m_y)_{ij} = 0.$$

Now, for γ_k , which is an arbitrary proper value of m_z , one will have:

$$[(m_x - i m_y) (m_x + i m_y)]_{kk} = \sum_l (m_x - i m_y)_{kl} (m_x + i m_y)_{lk},$$

in which all of the terms in the sum on the right-hand side will be zero except for possibly the term $\gamma_l = \gamma_k - 1$, if it exists. Hence, if $\gamma_k - 1$ is not a proper value of m_z then all of the terms in the sum will be zero, and one will have:

$$[(m_x - i m_y) (m_x + i m_y)]_{kk} = 0.$$

Now, one also has:

$$\begin{aligned} (m_x - i m_y) (m_x + i m_y) &= m_x^2 + m_y^2 + i [m_x, m_y] = m_x^2 + m_y^2 + m_z \\ &= m^2 - m_z^2 + m_z = m^2 + \frac{1}{4} - \left(m_z - \frac{1}{2}\right)^2. \end{aligned}$$

If $\gamma_k - 1$ is not a proper value of m_z then one will have:

$$\left[m^2 + \frac{1}{4} - \left(m_z - \frac{1}{2}\right)^2 \right]_{kk} = m^2 + \frac{1}{4} - \left(\gamma_k - \frac{1}{2}\right)^2 = 0,$$

so

$$\gamma_k = \frac{1}{2} \pm \theta, \quad \theta = \sqrt{m^2 + \frac{1}{4}}.$$

Briefly, if γ_k is a proper value of m_z then $\gamma_k - 1$, $\gamma_k - 2$, etc., ... will also be ones, until one arrives at a proper value that is equal to $\frac{1}{2} \pm \theta$. One will then obtain a decreasing sequence of proper values that will necessarily be bounded below, since all of the proper values of m_z are greater than $-m$. The sequence will then be:

$$\gamma_k, \gamma_k - 1, \gamma_k - 2, \dots, \theta - \frac{1}{2}.$$

The complete sequence of proper values is then:

$$-\theta + \frac{1}{2}, -\theta + \frac{3}{2}, \dots, \theta - \frac{3}{2}, \theta - \frac{1}{2},$$

and, in turn, it will be necessary that $\theta - \frac{1}{2} - (-\theta + \frac{1}{2}) = 2\theta - 1$ must be an integer; i.e., that one must have either $\theta = n$ or $\theta = (n + 1) / 2$ (n integer). The corresponding value of m^2 is $\theta^2 - \frac{1}{4}$, from the definition of θ . If we then set:

$$\theta = j + \frac{1}{2}$$

as the definition of j then j will either be a positive integer or zero (0, 1, 2, ...) or it will be a positive half-integer ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$), since m^2 is positive or zero. Since one will then have $\theta^2 - \frac{1}{4} = j(j+1)$, one will see that the possible proper values of m^2 have the form $j(j+1)$, with either $j = 0, 1, 2, \dots$, or $j = \frac{1}{2}, \frac{3}{2}, \dots$, and for the same given value of m^2 , m_z will have $2j+1$ possible proper values $-j, -j+1, \dots, j-1, j$.

In summary, the operator m^2 will have proper values $j(j+1)$ with $j = 0, 1, \dots$ or $j = \frac{1}{2}, \frac{3}{2}, \dots$, and for each given value of j , each of the operators m_x, m_y, m_z will have the $2j+1$ possible proper values $-j, -j+1, \dots, j-1, j$. Those are the conclusions that we can infer from the single fact that the operators m_x, m_y, m_z obey the commutation relations:

$$[m_x, m_y] = -i m_z, \quad [m_y, m_z] = -i m_x, \quad [m_z, m_x] = -i m_y.$$

If one applies those general results to the orbital kinetic moment then one will see that the operator M^2 in fact has the proper values $l(l+1)\frac{h^2}{4\pi^2}$, with $l = 0, 1, 2, \dots$, and for a given l , each of the three operators will, in fact, have the $2l+1$ possible proper values $-l\frac{h}{2\pi}, -(l-1)\frac{h}{2\pi}, \dots, (l-1)\frac{h}{2\pi}, l\frac{h}{2\pi}$, but now the number l cannot take on half-integer values $\frac{1}{2}, \frac{3}{2}, \dots$, as in the general case. That amounts to the fact that now m_x, m_y, m_z must satisfy not only the commutation relations (III.a), but also the more restrictive definitions:

$$m_x = i \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \quad m_y = i \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \quad m_z = i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

If one then considers the non-commutation relations (which are, as we have seen, coupled to the rotation group, which is the one that is most essential to the theory of kinetic moments) then one can think that if the orbital kinetic moments involve only the integer values of the number j of the general theory then other forms of the kinetic moments can involve the half-integer values. That is what the theory of spin presents.

4. SPIN

For quite some time, material particles have been considered to be characterized entirely by two constants: viz., their mass (or more precisely, their proper mass, in relativity theory) and their electric charge. However, the existence of various phenomena [e.g., anomalous Zeeman effect, fine structure of certain lines, etc. ⁽¹⁾] that are impossible to interpret by the quantum theories of the atom, even when one employs wave

⁽¹⁾ See L. de Broglie, *L'Électron Magnétique*, Hermann, Paris, 1934.

mechanics, have shown the necessity of attributing a third essential characteristic to electrons, in addition to their proper mass and charge, namely, their *spin*.

If one remains within the context of classical theories then the spin of the electron can be represented by a rotation of the electrified corpuscle around one of its diameters. Consequently, a proper kinetic moment, for which we reserve the name of *spin*, will exist for that corpuscle, as well as a proper magnetic moment that is due to the rotation of the charge. In order to interpret experimental facts, it will be necessary to attribute the value $\pm h / 2\pi$ to the proper kinetic moment of the electron, and the value $\frac{eh}{4\pi m_0 c}$ to its proper

magnetic moment, which is equal to one *Bohr magneton*; that is the *Uhlenbeck-Goudsmit hypothesis*. One then sees that for spin, one is led to make use of the possibility that is predicted by the general theory above that one can attribute a value for the kinetic moment that is equal to a half-integer times the unit $h / 2\pi$ (1 / 2 times that unit, for the electron).

More generally, it indeed seems that today there is good reason to attribute spin to any particle at the microscopic level. From the classical standpoint, that proper kinetic moment must be represented by a vector \mathbf{S} whose rectangular components are S_x, S_y, S_z , and whose length-squared will be:

$$S^2 = S_x^2 + S_y^2 + S_z^2.$$

From the quantum viewpoint, we must replace the classical quantities by operators, but *a priori* we do not know the form of those operators because we no longer know a classical expression that is capable of guiding our choice, as in the case of orbital kinetic moment. Meanwhile, what we can assume is that spin, which has the nature of a kinetic moment, must be coupled with the group of rotations in the same manner as the orbital kinetic moment.

If we introduce the quantum unit $h / 2\pi$ then we first write:

$$(S_x)_{\text{op}} = \frac{h}{2\pi} s_x, \quad (S_y)_{\text{op}} = \frac{h}{2\pi} s_y, \quad (S_z)_{\text{op}} = \frac{h}{2\pi} s_z,$$

and we will then assume that the operators s_x, s_y, s_z satisfy the non-commutation relations:

$$[s_x, s_y] = -i s_z, \quad [s_y, s_z] = -i s_x, \quad [s_z, s_x] = -i s_y,$$

which correspond to the relations that were assumed above for the m_x, m_y, m_z , and which express the relationship between spin and the rotation group here. Finally, in place of the classical quantity S^2 , we introduce a *total spin operator*:

$$(S^2)_{\text{op}} = \frac{h^2}{4\pi^2} s^2, \quad \text{with} \quad s^2 = s_x^2 + s_y^2 + s_z^2.$$

Under those conditions, it will result from the general proof that was given before that the possible proper values for S^2 will have the form $s(s+1) h^2 / 4\pi^2$, with either $s = 0, 1,$

2, ... or $s = \frac{1}{2}, \frac{3}{2}, \dots$, and that for a given value of s , each of the operators S_x, S_y, S_z will have the $(2s + 1)$ possible proper values:

$$-s \frac{h}{2\pi}, -(s+1) \frac{h}{2\pi}, \dots, (s-1) \frac{h}{2\pi}, s \frac{h}{2\pi}.$$

For the electron and other particles of spin $1/2$, one must take the operators in such a way that S^2 has the proper value $\frac{3}{4} \frac{h^2}{4\pi^2} = \frac{1}{2} \left(\frac{1}{2} + 1 \right) \frac{h^2}{4\pi^2}$ that corresponds to $s = 1/2$. S_x, S_y, S_z will then have the proper values $\pm \frac{1}{2} \frac{h}{2\pi}$, which indeed correspond to the spin of the electron, from the Uhlenbeck-Goudsmit hypothesis. We shall ultimately see how Dirac's theory specifies the form of the operators S_x, S_y, S_z .

Upon starting with the Dirac electron, which is considered to be a type of elementary particle, the process of "fusing" elementary corpuscles ⁽¹⁾ will permit us to construct particles that have several different states of total spin. For each of those states, the number s will have a well-defined value that can vary from $s = 0$ to $s = n$ for particles that are obtained fusing an even number $2n$ of constituents, and which can vary from:

$$s = \frac{1}{2} \quad \text{to} \quad s = \frac{2n+1}{2} = n + \frac{1}{2}$$

for particles that are composed of an odd number $2n + 1$ of constituents by fusion.

In this treatise, we shall confine ourselves to the theory of particles with $s = 1/2$, and the electron is that type of particle.

⁽¹⁾ See L. de Broglie, *Théorie Générale des particules à spin*, Gauthier-Villars, Paris, 1943.

CHAPTER IV

PROPER KINETIC MOMENTS FROM THE RELATIVISTIC VIEWPOINT

1. GENERALITIES

We shall soon see that Dirac's theory of the electron teaches us that in wave mechanics one must simultaneously introduce spin and relativity, which are closely linked with each other. That will then lead us to study how one might present the notion of kinetic moment from the relativistic viewpoint. The question might seem simple, but as we shall see, it is much more complicated than it would appear, notably for proper kinetic moments, or *spins*.

We assume that the general principles and formalism of special relativity are known. Meanwhile, we shall recall some points concerning the choice of variables. In the course of our presentation, we shall indeed sometimes employ complex Minkowski variables and sometimes real space-time variables, and in order to avoid confusion, it will be useful to specify the form and properties of those variables.

One can frame an event that takes place at a point at a certain instant in a Galilean reference system by giving the four world-coordinates (in the Minkowski sense) of that event. They are:

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict .$$

The fourth coordinate is pure imaginary.

With that choice of coordinates, the distance between two infinitely-close events in space-time is given in special relativity by a ds such that:

$$ds^2 = - \sum_j dx_j^2 .$$

ds^2 will then have a Euclidian form, and that is the essential advantage of Minkowski coordinates. There is then no reason to distinguish the covariant components of a tensor from its contravariant components, and one will have:

$$T_{ijk\dots}^{rst\dots} = T_{ijk\dots rst} .$$

In place of the Minkowski variables, one can also utilize real space-time coordinates (as is customary in general relativity), which are defined by:

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ct,$$

but one must adopt the pseudo-Euclidian expression:

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = dx_4^2 - dx_1^2 - dx_2^2 - dx_3^2 ,$$

which one can write, with a well-known notation, as:

$$ds^2 = \sum_{ik} g_{ik} dx^i dx^k,$$

with

$$g_{44} = +1, \quad g_{11} = g_{11} = g_{11} = -1, \quad g_{ik} = 0, \text{ for } i \neq k.$$

This ds^2 no longer has a Euclidian form, so it will be necessary to distinguish the covariant components from the contravariant components, and one will have:

$$T_{ijk\dots}^{rst\dots} = \sum_u g_{ut} T_{ijk\dots}^{st\dots u}.$$

Hence, from the values of the g_{ik} , in order to lower one of the indices 1, 2, 3, it is necessary to simply change the sign of the component, while in order to lower the index 4, no change needs to be made.

It is, moreover, easy to pass from world variables to the space-time variables, or conversely. If one knows the contravariant components $T^{ij\dots}$ of a tensor T when one employs the space-time variables then one will, in fact, get its components in the system of world variables by multiplying each T^{ij} by the factor $i = \sqrt{-1}$ as many times as the sequence $ij\dots$ contains the factor 4.

2. RELATIVISTIC REPRESENTATION OF ORBITAL KINETIC MOMENT

Suppose one has a corpuscle of mass m_0 . According to the classical conception, the sequence of its positions in space in the course of time is represented in space-time by its *world-line*. Each element of the world-line whose components are dx_i will satisfy the relation will satisfy the relation (in space-time coordinates):

$$ds = \sqrt{1 - \frac{dx^2 + dy^2 + dz^2}{dx_4^2}} dx_4 = c dt \sqrt{1 - \frac{v^2}{c^2}}.$$

v is the velocity of the corpuscle in the reference system employed at the instant t that corresponds to ds . At each point of the world-line, one can attach a quadri-vector that is defined by its covariant components:

$$u^i = \frac{dx^i}{ds}.$$

Upon multiplying this by the scalar invariant $m_0 c$, one will define the components of another quadri-vector – namely, the *impulse* \mathbf{p} of the corpuscle – whose components are:

$$p^i = m_0 c \frac{dx^i}{ds}.$$

In the Galilean system in which the corpuscle has a velocity v , one will have:

$$p^1 = p_x = \frac{m_0 v_x}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \dots, \quad \frac{W}{c} = p^4 = \frac{m_0 c}{\sqrt{1 - \frac{v^2}{c^2}}},$$

in which p_x, p_y, p_z are then components of the quantity of motion of the corpuscle, and W is its energy.

If we now define the quantities:

$$M^{ik} = x^i p^k - x^k p^i, \quad i, k = 1, 2, 3, 4$$

then we will get the components of an antisymmetric tensor of rank 2.

Explicitly, one will have:

$$\begin{aligned} M^{23} = -M^{32} &= y p_z - z p_y, & M^{14} = -M^{41} &= x \frac{W}{c} - ct p_x, \\ M^{31} = -M^{13} &= z p_y - y p_z, & M^{24} = -M^{42} &= y \frac{W}{c} - ct p_y, \\ M^{12} = -M^{21} &= x p_y - y p_x, & M^{34} = -M^{43} &= z \frac{W}{c} - ct p_z. \end{aligned}$$

The first three components behave like the components of a trivector under a spatial rotation. They define the orbital kinetic moment of the corpuscle *with respect to the origin of the coordinates*. The last three components, which also transform like the components of a trivector under a simple change of spatial coordinate axes, can be interpreted as the components of the barycentric moment with respect to the initial position (up to a factor). Indeed, upon introducing the mass of motion $m = m_0 / \sqrt{1 - v^2 / c^2}$, one will have:

$$M^{14} = mc (x - v_x t) = mc x_0, \quad M^{24} = mc (y - v_y t) = mc y_0, \quad M^{34} = mc (z - v_z t) = mc z_0$$

for uniform, rectilinear motion, in which x_0, y_0, z_0 are the coordinates of a corpuscle at the instant $t = 0$. If one makes a change of Galilean axes in space-time then the M^{ik} will transform like the components of an antisymmetric tensor of rank 2. For example, if one passes from a reference system $Oxyz$ to another system $Ox'y'z'$ with parallel axes that is in uniform, rectilinear motion with respect to the first one with a relative velocity βc in the Oz direction in such a way that the space-time coordinates submit to the simple Lorentz transformation:

$$x' = x, \quad y' = y, \quad z' = \frac{z - vt}{\sqrt{1 - \beta^2}}, \quad ct = \frac{ct - \beta z}{\sqrt{1 - \beta^2}},$$

then the components M'^{ik} of the tensor M in the second system will be expressed with the aid of the components M^{ik} of the same tensor in the first system by the classical formulas:

$$M'^{23} = \frac{M^{23} - \beta M^{14}}{\sqrt{1 - \beta^2}}, \quad M'^{31} = \frac{M^{31} + \beta M^{24}}{\sqrt{1 - \beta^2}}, \quad M'^{12} = M^{12},$$

$$M'^{14} = \frac{M^{14} + \beta M^{23}}{\sqrt{1 - \beta^2}}, \quad M'^{24} = \frac{M^{24} - \beta M^{31}}{\sqrt{1 - \beta^2}}, \quad M'^{34} = M^{34}.$$

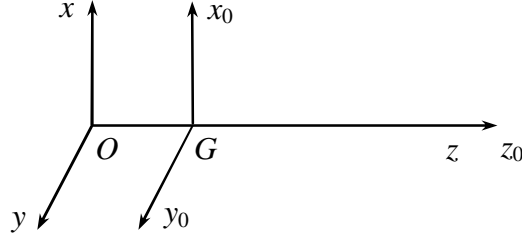
The results that we just recalled are often expressed by saying that one must consider any kinetic moment to be defined by an antisymmetric tensor of rank 2 (or at least, by the spatial components of such a tensor). In our opinion, that manner of speaking is a bit misleading. Indeed, in each reference system, the coordinate origin is an arbitrary point, and the kinetic moment with respect to that arbitrary point has no particular physical significance, in general. What does have physical significance is the kinetic moment with respect to a center that is endowed with physical properties, such a force center; i.e., the source particle of a central force field. Consider an observer that will pass that center with the velocity v of a hydrogen atom. In the proper system of the atom, the kinetic moment with respect to the origin, where one finds the nucleus, will have an important physical significance. However, if we transform the tensor \mathbf{M} that is provided in the proper system by the transformation above then the components M^{23} , M^{31} , M^{12} of the tensor in the fixed observer's system will give a kinetic moment with respect to its coordinate origin, which has no physical interest. Even for the fixed observer, what has the physical sense is the kinetic moment of the atomic electron around the moving nucleus and not the one with respect to an arbitrary coordinate origin. If one seeks to find how the fixed observer can represent the kinetic moment of the electron with respect to the comoving nucleus (which is, collectively, the proper kinetic moment of the moving H atom) then one will find that for each Galilean observer, there will exist a vector that represents that proper moment, but when one changes the Galilean observer, the components of that proper kinetic moment will not transform like the components of an antisymmetric tensor of rank 2. One then sees that the representation of a proper kinetic moment by an antisymmetric tensor of rank 2 is somewhat fallacious.

3. STUDY OF PROPER KINETIC MOMENT FROM THE RELATIVISTIC VIEWPOINT

Let there be a Galilean observer, whom we call "observer A"; he employs a Cartesian reference system $Oxyz$ and a time t . He then passes to a system that is defined by a particle M of proper mass m_0 that orbits around an attractive center G . We suppose that this system is animated with a uniform, rectilinear motion with respect to the Galilean observe A; i.e., we may link G with a Galilean reference frame $G x_0 y_0 z_0$. That hypothesis will raise some delicate questions that are related to the difficulty associated with making a *relativistic* definition of the center of gravity. Be that as it may, we shall assume the existence of a Galilean reference system $G x_0 y_0 z_0$ that accompanies the

system in its motion and is animated with respect to A with a uniform, rectilinear motion. When we say *proper kinetic moment* of the system, we mean the kinetic moment of the system with respect to G as it appears to the observer A .

In order to specify the definition of the proper kinetic moment, we first place ourselves in the reference system $G x_0 y_0 z_0$, which we call “the proper system.”



We suppose that the axes $Oxyz$ and $G x_0 y_0 z_0$ are parallel, and that the second one is animated with a velocity of βc with respect to the first one in the sense Oz , which is no loss of generality. The kinetic moment of the system will then be defined in the proper system by the spatial components of the antisymmetric tensor of rank 2:

$$M_{(0)}^{ik} = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} (x_{(0)}^i v_{(0)}^k - x_{(0)}^k v_{(0)}^i) = m_0 (x_{(0)}^i u_{(0)}^k - x_{(0)}^k u_{(0)}^i).$$

Explicitly, the kinetic moment will then be represented in the system $G x_0 y_0 z_0$ by a spatial vector $\mathbf{S}^{(0)}$ with the components:

$$(IV.a) \quad S_x^{(0)} = M_{(0)}^{23} = m_0 \frac{y_0 v_z^{(0)} - z_0 v_y^{(0)}}{\sqrt{1 - \frac{v^2}{c^2}}}, \dots$$

We now place ourselves with the observer in the system $Oxyz$. In that system, the antisymmetric tensor M will have the components:

$$M^{23} = \frac{M_{(0)}^{23} - \beta M_{(0)}^{24}}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad M^{31} = \frac{M_{(0)}^{31} + \beta M_{(0)}^{14}}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad M^{12} = M_{(0)}^{12},$$

(IV.b)

$$M^{14} = \frac{M_{(0)}^{14} + \beta M_{(0)}^{31}}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad M^{24} = \frac{M_{(0)}^{24} - \beta M_{(0)}^{23}}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad M^{34} = M_{(0)}^{34},$$

but the time origin is chosen in such a fashion that at time $t = 0$, the point G will coincide with O , so the components M^{23} , M^{31} , M^{12} will represent the components of the orbital

kinetic moment M of the particle with respect to the point O . Now, as we have remarked, that kinetic moment with respect to the origin of the coordinates has no physical interest. What does have physical interest for the observer A is the kinetic moment of the molecule that is moving with respect to its central point G , which is a quantity that is the proper kinetic moment of the comoving system for the observer A. How can the observer A mathematically define the spatial vector that represents that vector for him? For example, he might imagine some axes $G\xi\eta\zeta$ that are linked to G and coupled with $Oxyz$; those axes coincide with $Gx_0y_0z_0$. If I call them $G\xi\eta\zeta$ then one should recall that, at least as far as $G\xi$ is concerned, the lengths that A evaluates along that axis will differ, due to the Lorentz contraction, and similarly, the lengths that are evaluated in the proper system. The coordinates ξ, η, ζ of the molecule whose coordinates are x, y, z are:

$$\xi = x, \quad \eta = y, \quad \zeta = z - \beta ct$$

for A. Moreover, the components of the quantity of motion of the molecule with proper mass m_0 to the extent that those components provide the rotational motion around G are given by:

$$p_\xi = \frac{m_0 v_x}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad p_\eta = \frac{m_0 v_y}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad p_\zeta = \frac{m_0 (v_z - \beta c)}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Indeed, for the molecule whose total velocity is \mathbf{v} , $m_0 \beta c / \sqrt{1 - v^2/c^2}$ represents the part of the z -component of impulse that is due to the collective velocity βc . The spatial vector that represents the proper kinetic moment for the observer A will then be defined by the formulas:

$$S_x = m_0 \left[y \frac{v_z - \beta c}{\sqrt{1 - \frac{v^2}{c^2}}} - (z - \beta ct) \frac{v_y}{\sqrt{1 - \frac{v^2}{c^2}}} \right], \quad S_y = m_0 \left[(z - \beta ct) \frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}} - x \frac{v_z - \beta c}{\sqrt{1 - \frac{v^2}{c^2}}} \right],$$

(IV.c)

$$S_z = m_0 \left[x \frac{v_y}{\sqrt{1 - \frac{v^2}{c^2}}} - y \frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}} \right].$$

Upon comparing these with the expressions for $S_x^{(0)}$, $S_y^{(0)}$, $S_z^{(0)}$, one will see, first of all, that if the observer A is linked with the comoving system then he will find (as is obvious *a priori*) the vector $\mathbf{S}^{(0)}$ for the vector \mathbf{S} , because one will then have $\beta = 0$, $x = x_0$, ..., $v_z = v_z^{(0)}$. On the other hand, one will have the transformation formulas:

$$y = y_0, \quad x = x_0, \quad z - \beta ct = z_0 \sqrt{1 - \beta^2},$$

and the formulas for the component of velocity:

$$v_x = \frac{v_x^{(0)} \sqrt{1 - \beta^2}}{1 + \frac{\beta}{c} v_z^{(0)}}, \quad v_y = \frac{v_y^{(0)} \sqrt{1 - \beta^2}}{1 + \frac{\beta}{c} v_z^{(0)}}, \quad v_z = \frac{v_z^{(0)} + \beta c}{1 + \frac{\beta}{c} v_z^{(0)}}.$$

One first infers that:

$$\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{v_0^2}{c_0^2}}} \cdot \frac{1 + \frac{\beta}{c} v_z^{(0)}}{\sqrt{1 - \beta^2}},$$

and then:

$$\frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{v_x^{(0)}}{\sqrt{1 - \frac{v_0^2}{c_0^2}}}, \quad \frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{v_x^{(0)}}{\sqrt{1 - \frac{v_0^2}{c_0^2}}}, \quad \frac{v_z - \beta c}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{v_z^{(0)} \sqrt{1 - \beta^2}}{\sqrt{1 - \frac{v_0^2}{c_0^2}}}.$$

Upon comparing the expressions (IV.a) and (IV.c), one will easily infer that:

$$S_x = S_x^{(0)} \sqrt{1 - \beta^2}, \quad S_y = S_y^{(0)} \sqrt{1 - \beta^2}, \quad S_z = S_z^{(0)}.$$

Those transformation formulas are completely different from the ones for the spatial components of an antisymmetric tensor of rank 2. Thus, for example, if one lets β tend to 1 then, from formulas (IV.b), the spatial vector whose components are M^{23} , M^{31} , and M^{12} will tend to be perpendicular to Oz , while \mathbf{S} will tend to lie along Oz .

Indeed, for a given observer A, one can find an antisymmetric tensor that has the spatial components S_x, S_y, S_z in the system $Oxyz$ and $S_x^{(0)}, S_y^{(0)}, S_z^{(0)}$ in the proper system. Indeed, in order for that to be true, it will suffice to define that tensor by the formulas:

$$\begin{aligned} S^{23} &= S_x, & S^{31} &= S_y, & S^{12} &= S_z, \\ S^{14} &= 0, & S^{24} &= 0, & S^{34} &= 0, \end{aligned}$$

which will indeed give:

$$(IV.d) \quad S_{(0)}^{23} = S_x^{(0)}, \quad S_{(0)}^{31} = S_y^{(0)}, \quad S_{(0)}^{12} = S_z^{(0)},$$

from the transformation formulas for antisymmetric tensors. However – and this is essential – the tensor thus-defined will change when one passes from one Galilean system

A to another one A' that is in relative motion with respect to A. In order to see that, it will suffice to remark that if one has just one tensor \mathbf{S} then that tensor must have its S^{i4} components equal to zero in any Galilean system, which is impossible. We shall recover that conclusion later on by a different method.

Finally, we could define a spatial vector \mathbf{S} for each Galilean observer that defines the proper kinetic moment of the system considered, but we could not attach that vector to a mathematical entity with a tensorial character in space-time in a unique fashion.

If one seeks to define the tensor \mathbf{S} in such a fashion that one has

$$S_{(0)}^{23} = S_x^{(0)}, \quad S_{(0)}^{31} = S_y^{(0)}, \quad S_{(0)}^{12} = S_z^{(0)}, \quad S_{(0)}^{14} = S_{(0)}^{24} = S_{(0)}^{34} = 0$$

in the proper system then the first three components of that tensor in another Galilean system will be:

$$S^{23} = \frac{S_x^{(0)}}{\sqrt{1-\beta^2}}, \quad S^{31} = \frac{S_y^{(0)}}{\sqrt{1-\beta^2}}, \quad S^{12} = \frac{S_z^{(0)}}{\sqrt{1-\beta^2}},$$

and that will not be the three components of the proper moment in that system. That remark will prove useful to us in our study of von Weysenhoff's theory.

In wave mechanics, one always seeks to associate any particle with its densities that have the character of field magnitudes. When we seek to define the proper kinetic moment of a particle (e.g., electron spin) in wave mechanics, it would be natural to define the spin by a vector \mathbf{S} of the form:

$$\mathbf{S} = \int_D \boldsymbol{\sigma} d\tau,$$

in which $\boldsymbol{\sigma}$ will be the *density* of proper moment.

In the proper system, we will then have:

$$S_x^{(0)} = \int_D \sigma_x^{(0)} d\tau, \quad S_y^{(0)} = \int_D \sigma_y^{(0)} d\tau, \quad S_z^{(0)} = \int_D \sigma_z^{(0)} d\tau,$$

and we will suppose that $\boldsymbol{\sigma}^{(0)}$ is a function of x_0, y_0, z_0 , but independent of t_0 . If we would like to perform the integrations upon appealing to the variables x, y, z of an observer A then we must replace $d\tau$ with $d\tau / \sqrt{1-\beta^2}$, due to the Lorentz contraction, and we will find that:

$$S_x^{(0)} = \frac{S_x}{\sqrt{1-\beta^2}} = \int_D \frac{\sigma_x^{(0)}}{\sqrt{1-\beta^2}} d\tau, \quad S_y^{(0)} = \frac{S_y}{\sqrt{1-\beta^2}} = \int_D \frac{\sigma_y^{(0)}}{\sqrt{1-\beta^2}} d\tau,$$

$$S_z^{(0)} = S_z = \int_D \frac{\sigma_z^{(0)}}{\sqrt{1-\beta^2}} d\tau,$$

so one will infer:

$$\sigma_x = \sigma_x^{(0)}, \quad \sigma_y = \sigma_y^{(0)}, \quad \sigma_z = \frac{\sigma_z^{(0)}}{\sqrt{1-\beta^2}}.$$

The quantities σ_x , σ_y , σ_z then transform like the three rectangular spatial components of a quadri-vector Σ whose fourth component $\sigma_4^{(0)}$ will be zero in the proper system. When β tends to 1, the vector σ will lie along the direction of motion. Thus, whereas the proper kinetic moment does not have a well-defined tensorial character, by contrast, one can define a *proper kinetic moment density* with the aid of a quadri-vector Σ whose temporal component is zero in the proper system (which seems natural from the physical viewpoint). Moreover, the latter condition permits one to express σ_4 as a function of the spatial vector σ in no particular Galilean system. Indeed, the scalar product:

$$(\Sigma \cdot ds) = \sigma_4 dt - \sigma_x dx - \sigma_y dy - \sigma_z dz$$

is zero, as one sees upon placing oneself in the proper system. In other words, the quadri-vector Σ is orthogonal to the world-line of the particle in space-time. One will then have (since $\Sigma \cdot ds$ is an invariant):

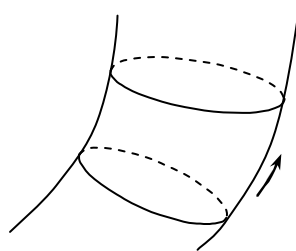
$$\sigma_4 = \frac{1}{c} (\sigma \cdot \mathbf{v})$$

in any Galilean system. We shall take up the problem of the relativistic representation of the proper kinetic moment once more after we have taken our inspiration from the work of Olivier Costa de Beauregard, and we will recover the same results in a different way.

4. GENERAL-RELATIVISTIC THEORY OF PROPER KINETIC MOMENTS

Consider a fluid in motion that is conceived in the classical fashion. The set of world-lines and its various elements form a world-tube that occupies a four-dimensional domain in space-time that extends in the time direction.

Cut that tube with three-dimensional endcaps (hyper-endcaps) C, C', \dots . Suppose that our fluid is endowed with a proper kinetic moment, such that each fluid element transports its moment, and with Costa de Beauregard, attempt to define a *density* of proper kinetic moment at each point of the world-tube, which, when integrated over an arbitrary endcap C , will give a proper kinetic moment, or more precisely, an antisymmetric second-rank tensor whose spatial components will define a kinetic moment. Since the endcap C is three-dimensional, an element of that endcap can be defined to be a parallelepiped that is constructed from three small vectors ds^I, ds^{II}, ds^{III} that are contained in that multiplicity:



As one knows, that volume element can be considered to be a quadri-vector, or more precisely, a completely-antisymmetric third-rank tensor whose projection onto the index 1 (i.e., the projection of the volume onto the hyperplane that is perpendicular to the x -axis) is given by the determinant:

$$[dx_j, dx_k, dx_l] = \begin{vmatrix} dx_j^I & dx_k^I & dx_l^I \\ dx_j^{II} & dx_k^{II} & dx_l^{II} \\ dx_j^{III} & dx_k^{III} & dx_l^{III} \end{vmatrix}.$$

It is then natural to try, as Costa de Beauregard did, to write the proper kinetic moment that is attached to the endcap C in the form:

$$dS_{ij} = \sum_{k=1}^4 \sigma^k [dx_j, dx_k, dx_l],$$

in which the four σ^k form the components of a space-time quadri-vector (or rather – what amounts to practically the same thing in special relativity – a completely-antisymmetric third-rank tensor). That quadri-vector is the “proper kinetic moment density” that is defined at any point of the fluid.

Upon integrating dS_{ij} over the chosen endcap C , one will obtain the quantity:

$$S_{ij}^{(0)} = \int_C \sum_{k=1}^4 \sigma^k [dx_j, dx_k, dx_l].$$

We have written $S_{ij}^{(0)}$, because we do not further know whether that quantity does or does not depend upon the chosen endcap C . In order to examine that question, it is interesting to compare the preceding definitions with the ones that are usually adopted in special relativity for electric charge and current.

In relativity, one defines the motion of a distribution of electricity by a space-time quadri-vector \mathbf{J} whose spatial components j_x, j_y, j_z are equal to the components of the current density, and whose time component is equal to $c\rho$; i.e., to c times the density of electricity. The collective motion of the distribution of electricity in the course of time will be represented by a “world-tube” that is composed of the set of world-lines over its various elements. On the walls of the tube, the quadri-vector \mathbf{J} is, by definition, always tangent to the wall, which expresses the physical fact that the electricity does not cross that wall. We can further cut the world-tube by a three-dimensional endcap C .

With the same notations as above, we define the quantity:

$$\begin{aligned} F_{(C)}[\mathbf{J}] &= \int_C (\rho c [dx_1, dx_1, dx_1] - j_3 [dx_4, dx_1, dx_2] - j_2 [dx_3, dx_4, dx_1] - j_1 [dx_2, dx_3, dx_4]) \\ &= \int_C \begin{vmatrix} dx_1^I & dx_2^I & dx_3^I & dx_4^I \\ dx_1^{II} & dx_2^{II} & dx_3^{II} & dx_4^{II} \\ dx_1^{III} & dx_1^{III} & dx_1^{III} & dx_1^{III} \\ j_1 & j_2 & j_3 & \rho c \end{vmatrix}. \end{aligned}$$

The quantity $F_{(C)}[\mathbf{J}]$ is the flux of the quadri-vector \mathbf{J} that crosses the endcap C . It possesses two essential properties:

1. $F_{(C)}[\mathbf{J}]$ is a relativistic pseudo-invariant; i.e., for a given endcap C , it will have the same value, no matter what reference system is used to evaluate it.

That property results from the tensorial character of \mathbf{J} and the hypersurface elements; it is entirely independent of the conservation of electricity.

2. If one considers two different endcaps C and C' that cut the world-tube then one will have:

$$F_{(C)}[\mathbf{J}] = F_{(C')}[\mathbf{J}].$$

That second property expresses the conservation of electricity. Indeed, when one takes into account the positive sense that is chosen along the tube, the flux-divergence theorem will apply to the space-time domain D that is composed of the interior of the world-tube between the given endcaps C and C' :

$$\begin{aligned} F_{(C)}[\mathbf{J}] = F_{(C')}[\mathbf{J}] &= \int_D \left(\frac{\partial \rho}{\partial t} + \text{div } \mathbf{j} \right) [dx_1, dx_2, dx_3, dx_4] \\ &= \int_D \text{Div } \mathbf{J} [dx_1, dx_2, dx_3, dx_4], \end{aligned}$$

in which:

$$\text{Div } \mathbf{J} = \frac{\partial \rho}{\partial t} + \text{div } \mathbf{j}$$

is the quadri-dimensional divergence of \mathbf{J} . If one assumes the conservation of electricity, as expressed by the equation $\text{Div } \mathbf{J} = 0$, then one will have:

$$F_{(C')}[\mathbf{J}] = F_{(C)}[\mathbf{J}].$$

If the quadri-vector \mathbf{J} does not obey the continuity equation then the first of the properties of $F_{(C)}[\mathbf{J}]$ – namely, its invariance – will persist, but the second one – namely, its constancy when one displaces the endcap – will not. That indeed sheds light upon the difference between the invariance of electric charge and its conservation.

Briefly, the quantity $F_{(C)}[\mathbf{J}]$ will have the same value for any endcap C that cuts the world-tube, so that value will be characteristic of the tube, as a whole. By definition, one considers it to measure the “total electric charge” (up to a factor $1/c$) of the distribution. In order to see that the definition conforms to the usual notion of charge, it is sufficient to remark that an observer will naturally take an endcap C that is composed of space-time points that are simultaneous for him, and that he will define the electric charge by:

$$e = \int_V \rho dv,$$

in which V is the volume that the distribution of electricity at the proper time instant t when the integration is performed.

After that rapid study of the relativistic definition of electric charge, we return to the definition that was given above for the proper kinetic moment:

$$S_{ij}^{(C)} = \int_C \sum_{k=1}^4 \sigma^k [dx_i, dx_j, dx_k].$$

For a given value of the indices i and j , there will be only two non-zero terms in the sum on the right-hand side (due to the antisymmetry of the bracket). However – and this is an essential point – the second-rank antisymmetric tensor thus-defined will depend upon the choice of the endcap C .

Physically, it is natural (and we will even see that it is almost necessary in the quantum theory of spin) to define a tensor \mathbf{S} in each Galilean system with the aid of an endcap C whose points are all simultaneous in a reference system. If the motion of the fluid is a collective motion of a permanent character then the tensor thus-defined for a certain Galilean observer will remain constant in the course of time, but when one passes from the first observer to the second one in relative motion with respect to the first one, one will pass from an endcap C_1 to an endcap C_2 , and as a result, from a tensor \mathbf{S}_1 to a tensor \mathbf{S}_2 . The tensors \mathbf{S} are then defined *with respect to an observer*. If all of the world-lines of the fluid elements are parallel then one can consider the endcaps C that are orthogonal to the world-lines in space-time. They correspond to the fluid volume for an observer that is coupled with its motion (proper system). One can define a tensor \mathbf{S} by appealing to its endcaps; i.e., by placing oneself in the proper system. We shall see that this is what van Weyssenhoff did. However, the tensor thus-obtained will have the components of the proper kinetic moment for its spatial components only *in the proper system*. The same tensor, when envisioned in a Galilean system other than the proper system, will no longer have the components of the proper kinetic moment for its spatial components.

For each observer, if we integrate in the proper space of that observer then we will get:

$$S_{ij} = \int_C \sigma^k [dx_i, dx_j, dx_k], \quad i, j, k = 1, 2, 3,$$

$$S_{i4} = 0, \quad i = 1, 2, 3.$$

We have previously encountered a tensor of that type.

For each observer, we will then get a tensor $\mathbf{S}^{(1)}$ that is linked with the endcap C that defines the “space” of that observer.

Naturally, a second observer can, in principle, calculate the components $S_{ij}^{(1)}$ of the tensor $\mathbf{S}^{(1)}$ of the first observer, but they are of no interest to us, and we will employ the tensor $\mathbf{S}^{(2)}$ that corresponds to it for that second observer, and whose components S_{i4} are zero. As we have pointed out already, the fact that the $i4$ components are zero will suffice to show that the various tensors $\mathbf{S}^{(1)}$, $\mathbf{S}^{(2)}$, ... are different from each other, because a second-rank antisymmetric tensor cannot have its S_{i4} equal to zero in every Galilean system.

The preceding remarks will find their place in the theory of electron spin. The propagation of the wave ψ of a spinning particle under consideration will define a world-tube in space-time, and we will be led to consider a “spin density” that is defined by a quadri-vector whose four components will be given by formulas of the type:

$$\sigma^i = \psi^* (\sigma^i)_{\text{op}} \psi,$$

in which $(\sigma^i)_{\text{op}}$ is an operator whose form we shall have to specify. The preceding definition conforms to the general definition of the densities in wave mechanics. According to those same principles, one can calculate the value of the tensor \mathbf{S} at each instant t for each observer by integrating over the endcap C of the world-tube that is formed by the space-time points that are *simultaneous* for the observer that is envisioned at his instant t . If the motion of the particle is permanent then the tensor \mathbf{S} will remain the same in the course of time for each observer, but even in that case, it will change when one passes from one Galilean observer to another. That is the essential point, because it will result from it that the mean values:

$$S_i = S_{jk} = \int_D \sigma^i d\tau = \int_D \psi^* (\sigma^i)_{\text{op}} \psi d\tau$$

(in which i, j, k form an *even* permutation of the indices 1, 2, 3) will not transform like the spatial components of a second-rank antisymmetric tensor when one passes from one Galilean observer to another. One must be very considerate of that situation, which is often misunderstood, when one tries to assimilate spin to an antisymmetric space-time tensor.

In reality, when one is occupied with relativistic variances, it will be preferable to consider only the “spin density” which has a completely well-defined tensorial character, since the σ^i that are defined at each point of space-time transform like the components of a quadri-vector. The integral spin must be envisioned as a well-defined spatial vector in each reference system rather than as a tensor. However, as we will see, it is a vector without variance that is defined in space-time that has a physical sense in quantum theory. That is an example of the very frequent contradictions between quantum and relativistic ideas that present themselves.

5. RELATIVISTIC ASPECT OF PROPER MAGNETIC MOMENTS

We have seen that spinning particles also have a proper magnetic moment. An examination of that magnetic moment will lead to some conclusions that are quite analogous to the one in the preceding paragraphs. It is not the magnetic moment itself that has a well-defined tensorial variance, but the magnetic moment density (magnetization). More precisely, one cannot separate the relativistic study of the proper magnetic moment from that of the proper electric moment, since the two notions are coupled to each other from the relativistic viewpoint like those of the electric field and the magnetic field. Just as the six components of the electric and magnetic fields are

united in a second-rank antisymmetric tensor in relativistic electromagnetism, the *densities* of electric and magnetic moments will similarly form a second-rank antisymmetric tensor. Let \mathbf{M} be the spatial vector of “magnetic moment” and let \mathbf{P} be the spatial vector of “dielectric moment,” so the corresponding densities $\boldsymbol{\mu}$ and $\boldsymbol{\pi}$ are such that, by definition:

$$M_i = \int_D \mu_i d\tau, \quad P_i = \int_D \pi_i d\tau, \quad i = 1, 2, 3.$$

Upon setting:

$$\begin{aligned} m_{23} = -m_{32} = \mu_x = \mu_1, & \quad m_{31} = -m_{13} = \mu_y = \mu_2, & \quad m_{12} = -m_{21} = \mu_z = \mu_3, \\ m_{14} = -m_{41} = \pi_x = \pi_1, & \quad m_{24} = -m_{42} = \pi_y = \pi_2, & \quad m_{34} = -m_{43} = \pi_z = \pi_3, \end{aligned}$$

one will define a second-rank antisymmetric tensor \mathbf{m} . From the transformation formulas for tensors of that type, one will see that if, for example, the electric moment of a body in a certain Galilean system in which it is at rest is found to be zero without its magnetic moment also being zero then in another system that is in motion with respect to the first one, the body will possess both a magnetic moment and an electric moment. One can say that upon passing from the first system to the second one, one will see that the body that is magnetized in its proper system without being electrically polarized will become electrically polarized as a result of its motion. That is precisely what happens for the electron. When envisioned in a system in which it is at rest, the electron will have a proper magnetic moment, but not a proper electric moment; when envisioned in a system in which it is in motion, it will have both a proper magnetic moment and a proper electric moment.

It is easy to find the expression for the vector $\boldsymbol{\pi}$ as a function of the vector $\boldsymbol{\mu}$ for a magnetic particle in uniform motion. Indeed, in the proper system of the particle, one will have $\boldsymbol{\pi}^{(0)} = 0$, by hypothesis. Now, since the tensor \mathbf{m} is antisymmetric, one will have:

$$(IV.e) \quad \sum_{l=1}^3 m_{il} dx_l - m_{i4} dx_4 = 0,$$

as one will easily see upon writing that invariant in the proper system. If one sets $i = 1, 2, 3$ in that equation then one will find that:

$$m_{ij} dx_j + m_{ik} dx_k = m_{i4} c dt,$$

in which i, j, k form an even permutation of the indices 1, 2, 3. Taking into account the values of m_{ik} , one will find that:

$$\boldsymbol{\pi} = \frac{1}{c} [\mathbf{m} \times \mathbf{v}],$$

which will give the expression for $\boldsymbol{\pi}$ in any Galilean system.

Moreover, if one sets $i = 4$ in (IV.e) then one will get:

$$(\boldsymbol{\pi} \cdot \nabla) = 0.$$

The vector $\boldsymbol{\pi}$ that is generated by the motion will always be normal to the velocity \mathbf{v} then. Those formulas are valid in Dirac's theory of the electron, and more generally, for any particle with a proper magnetic moment.

6. RELATIONSHIP BETWEEN MAGNETO-ELECTRIC MOMENT AND SPIN

The fundamental idea of the Uhlenbeck-Goudsmit hypothesis is that the electron possess a proper angular momentum that is equal to $\frac{1}{2} \frac{h}{2\pi} = \frac{h}{4\pi}$ and a proper magnetic moment that is collinear with it and has a value of $\frac{eh}{4\pi m_0 c}$ (viz., a Bohr magneton). Those two moments are then represented by vectors with the same direction and opposite senses (due to the *negative* charge of the electron), so the proper magnetic moment is deduced from the proper kinetic moment – or spin – by multiplying by the factor $\frac{-eh}{m_0 c}$:

$$\begin{array}{ccc} \boldsymbol{\mathfrak{M}} & \frac{h}{4\pi} & \\ \longleftarrow & \text{---} & \longrightarrow \\ \frac{-eh}{4\pi m_0 c} & & \mathbf{M} \end{array}$$

The considerations that we just developed permit us to see that this relation between the moments is exact only in the *proper* system of the particle. Indeed, in any Galilean system, the proper kinetic moment or spin is represented by a vector with components.

$$S_x = \int \sigma_1 d\tau, \quad S_y = \int \sigma_2 d\tau, \quad S_z = \int \sigma_3 d\tau,$$

in which $\sigma_x, \sigma_y, \sigma_z$ are the spatial components of the “spin density” quadri-vector that was considered above. The proper magnetic moment of the particle and the proper electric moment that it is associated with as a result of its motion are represented by two vectors with components:

$$\begin{aligned} \mathfrak{M}_x &= \int \mu_x d\tau = \int m_{23} d\tau, \dots \\ \mathfrak{P}_x &= \int \pi_x d\tau = \int m_{14} d\tau, \dots, \end{aligned}$$

in which m_{23}, \dots are the components of the second-rank antisymmetric tensor of the “density of magnetic and electric moment” that was defined before. When one places oneself in the proper system of the particle, one will have:

$$\sigma_4 = 0, \quad m^{14} = m^{24} = m^{34} = 0,$$

and moreover:

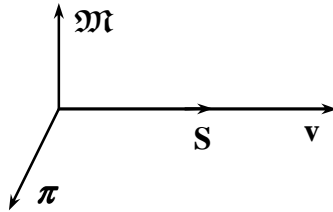
$$m_{23} = \frac{-e}{m_0 c} \sigma_1, \quad m_{31} = \frac{-e}{m_0 c} \sigma_2, \quad m_{12} = \frac{-e}{m_0 c} \sigma_3,$$

so:

$$\mathfrak{P} = 0, \quad \frac{\mathfrak{M}}{S} = -\frac{e}{m_0 c}.$$

The Uhlenbeck-Goudsmit hypothesis is then verified in the proper system.

However, that will no longer be true in another Galilean system in which the particle is in motion. From the formulas for the transformation of the components of $\boldsymbol{\sigma}$ and \mathbf{m} , one will see that the more rapid the motion of the particle, the more rapidly that the vector \mathbf{S} will tend lie along the trajectory, and the vector \mathfrak{M} will become normal to it, since the vector $\boldsymbol{\pi}$ will always be normal to both \mathfrak{M} and the velocity \mathbf{v} , by reason of the formulas $\boldsymbol{\pi} = \frac{1}{c} [\boldsymbol{\mu} \cdot \mathbf{v}]$ and $(\boldsymbol{\pi} \cdot \mathbf{v}) = 0$. In the limit as v tends to c , the three vectors will take on the following disposition:



which recalls the disposition of the electromagnetic vectors in an electromagnetic wave, which is a remark that served to guide me in the elaboration of my theory of the photon.

The proper magnetic moments and proper kinetic moment vectors are then collinear (in the opposite sense) only in the proper system of the electron.

7. JAN VAN WEYSSENHOFF'S THEORY

Jan van Weyssenhoff published a series of articles ⁽¹⁾ in which, developing some ideas of Mathisson, Lubanski, and Frenkel, he established a relativistic dynamics of spinning fluids and spinning particles. An attempt along those lines was also made by Olivier Costa de Beauregard in France.

I would like to summarize briefly some of van Weyssenhoff's calculations here, without entering into the details or pursuing the developments, in order to show the relationships with the concepts that were presented above.

The author started with the idea that one must introduce a second-rank antisymmetric tensor into the classical theory of spin that represents the *density of proper kinetic moment*. What we said before shows that we must express some reservations in regard to that hypothesis, since in reality, the spin density is represented by a quadri-vector, and not an antisymmetric tensor.

⁽¹⁾ Acta Physica Polonica 9 (1947), 8-55.

Weyssenhoff let s^{ik} denote the components of the tensor that he introduced and defined the two three-dimensional vectors: \mathbf{s} , with components s^{23} , s^{31} , s^{12} , and \mathbf{q} , with components s^{14} , s^{24} , s^{34} . The condition he posed was:

$$(IV.f) \quad s_{\alpha\beta} u^\beta = s^{\alpha\beta} u_\beta = 0$$

(with a summation over β), in which the u^β are the components of the world-velocity:

$$u^1 = \frac{v_x}{\sqrt{1-\beta^2}}, \quad u^2 = \frac{v_y}{\sqrt{1-\beta^2}}, \quad u^3 = \frac{v_z}{\sqrt{1-\beta^2}}, \quad u^4 = \frac{ic}{\sqrt{1-\beta^2}},$$

from which it will result that the components s^{14} , s^{24} , s^{34} of \mathbf{s} will be annulled in the proper system. The relation $s^{\alpha\beta} u_\beta = 0$ is equivalent to:

$$\mathbf{q} = \frac{1}{c} (\mathbf{s} \times \mathbf{v}) \quad \text{and} \quad (\mathbf{q} \cdot \mathbf{s}) = 0.$$

Weyssenhoff's tensor $s^{\alpha\beta}$ indeed represents (up to a constant factor $-e / m_0 c$) the tensor m^{ik} (density of proper electric and magnetic moment), but from what we said above, it will truly represent the proper kinetic moment only in the proper system, in which one will have:

$$s_{(0)}^{23} = \sigma_{(0)}^1, \quad s_{(0)}^{31} = \sigma_{(0)}^2, \quad s_{(0)}^{12} = \sigma_{(0)}^3, \quad s^{14} = s^{24} = s^{34} = 0,$$

so, upon integrating over the proper volume:

$$S_{(0)}^{23} = S_{(0)}^1 = S_x^{(0)}, \dots, \quad \text{with} \quad S_{(0)}^{23} = \int s_{(0)}^{23} d\tau_0, \dots,$$

which is precisely formulas (IV.d).

However, in a reference system other than the proper system, the tensor $s^{\alpha\beta}$ will no longer have a simple relationship to the spin, which is always defined by $S_x = \int \sigma^1 d\tau$, ..., because the three quantities $\int s^{23} d\tau = S^{23}$, $\int s^{31} d\tau = S^{31}$, and $\int s^{12} d\tau = S^{12}$ will no longer coincide with S_x , S_y , S_z . In particular, in a system in which $v \sim c$, the vector $\mathbf{S} = (S_x, S_y, S_z)$ lies along the velocity, while the vector S^{23}, S^{31}, S^{12} is perpendicular to it.

Weyssenhoff wrote the conservation of energy and impulse in the form:

$$\partial_\beta T^{\alpha\beta} = 0,$$

in which $T^{\alpha\beta}$ defines the energy-impulse tensor. In the case of spinning fluids, one sets $T^{\alpha\beta} = \mu_0 u^\alpha u^\beta$, in which μ_0 is the mass density. The tensor $T^{\alpha\beta}$ will then be symmetric. Weyssenhoff, adopting the viewpoint of Costa de Beauregard, who stressed its importance several times, did not suppose that the tensor $T^{\alpha\beta}$ was symmetric, and wrote:

$$T^{\alpha\beta} = \mathcal{G}^\alpha u^\beta,$$

in which \mathcal{G}^α is the quadri-vector of *proper linear impulse density*. We do not assume that \mathcal{G}^α is collinear with u^α , as one does in the absence of spin.

With Weyssenhoff, we now let Greek indices denote world-indices that run from 1 to 4, and let Latin indices denote ordinary spatial indices that run from 1 to 3. Let g^i denote the components of the quantity of motion in the ordinary sense; we will have:

$$T^{\alpha\beta} = \left\| \begin{array}{c|c} g^i v^k & c \mathbf{g} \\ \hline c \mu \mathbf{v} & c^2 \mu \end{array} \right\| = \begin{vmatrix} g^1 v^1 & g^1 v^2 & g^1 v^3 & c g^1 \\ g^2 v^1 & g^2 v^2 & g^2 v^3 & c g^2 \\ g^3 v^1 & g^3 v^2 & g^3 v^3 & c g^3 \\ c \mu v_1 & c \mu v_2 & c \mu v_3 & c^2 \mu \end{vmatrix}.$$

The relation $\partial_\beta T^{\alpha\beta} = 0$ expresses the conservation of energy and impulse. The relation $\partial_\beta T^{\alpha\beta} = 0$, which is equivalent to the preceding one in the case of the absence of spin, will no longer be exact here. The quantity of motion \mathbf{g} and the velocity \mathbf{v} will no longer be collinear. In the proper system of the particle, the energy is proportional to the rest mass.

Weyssenhoff then introduced two types of derivative with respect to time: First, the classical Lagrangian derivative (which follows the particle):

$$d_t f = \partial_t f + v^k \partial_k f,$$

and the derivative (for densities):

$$D_t f = d_t f + f \partial_k v^k = \partial_t f + \partial_k (f v^k).$$

One has:

$$d_t (f d\tau) = (D_t f) d\tau, \quad d_t \int f d\tau = \int (D_t f) d\tau,$$

because:

$$d_t (d\tau) = (\text{div } \mathbf{v}) d\tau.$$

By analogy, one defines derivation with respect to proper time along the world-line of a particle in Minkowski space:

$$\begin{aligned} d_{t_0} f &= \dot{f} = u^\nu d_\nu f, \\ D_{t_0} f &= d_{t_0} f + f \partial_\nu u^\nu = \partial_\nu (f u^\nu). \end{aligned}$$

If $d\Omega$ denotes a four-dimensional volume element then one will have:

$$d_{t_0} (f d\Omega) = (D_{t_0} f) d\Omega,$$

since:

$$d_{t_0} (d\Omega) = d\Omega \text{Div } \mathbf{u} = d\Omega \partial_\nu u^\nu.$$

However, $d\Omega = d\bar{\tau}_0 \cdot dt_0$, so:

$$d_{t_0} (f d\tau_0) = (D_{t_0} f) d\tau_0 .$$

One will then have:

$$\partial_\beta T^{\alpha\beta} = \partial_\beta (\mathcal{G}^\alpha u^\beta) = D_{t_0} \mathcal{G}^\alpha ,$$

so:

$$D_{t_0} \mathcal{G}^\alpha = 0 ,$$

which expresses the conservation of impulse.

Weysenhoff expressed the conservation of total moment of the quantity of motion (orbital + spin) by writing the condition:

$$D_{t_0} (x^\alpha \mathcal{G}^\beta - x^\beta \mathcal{G}^\alpha) + D_{t_0} s^{\alpha\beta} = 0 .$$

That expresses the conservation in the proper system; furthermore, $s^{\alpha\beta}$ represents spin only in that system.

From the definitions that were assumed, one will have:

$$D_{t_0} (f g) = f D_{t_0} g + g d_{t_0} f = g D_{t_0} f + f d_{t_0} g = d_{t_0} f + d_{t_0} g \partial_\nu u^\nu$$

for arbitrary f and g . The conservation equation will then be written:

$$D_{t_0} s^{\alpha\beta} = \mathcal{G}^\alpha u^\beta - \mathcal{G}^\beta u^\alpha = T^{\alpha\beta} - T^{\beta\alpha} ,$$

One sees that the existence of spin is related to the asymmetry of the tensor T , as Costa de Beauregard has observed. Multiply the preceding equation by u_β and note that $u^\beta u_\beta = -c^2$. Upon setting:

$$\mu_0 = -\frac{1}{c^2} u_\beta \mathcal{G}^\beta ,$$

one will get:

$$\mathcal{G}^\alpha = \mu_0 u^\alpha - \frac{1}{c^2} u_\beta D_{t_0} s^{\alpha\beta} = \mu_0 u^\alpha + \frac{1}{c^2} \dot{s}^{\alpha\beta} u_\beta ,$$

in which the last expression is obtained by applying the operator D_{t_0} to the equation $s^{\alpha\beta} u_\beta = 0$, and setting:

$$\dot{u}_\beta = d_{t_0} u_\beta .$$

Having thus defined the densities of a tensorial character, the author integrates them in order to obtain integral quantities. However, that operation will leave the variances intact only if one integrates *over proper volumes*. One then sets:

$$\begin{aligned} G^\alpha &= \int \mathcal{G}^\alpha d\tau_0 , \\ S^{\alpha\beta} &= \int s^{\alpha\beta} d\tau_0 , \end{aligned}$$

$$m_0 = \int \mu_0 d\tau_0.$$

The third definition is obvious, but there are some interesting remarks to be made about the second one. In the proper system, the tensor $S^{\alpha\beta}$ will have spatial components $S_{(0)}^{23}, S_{(0)}^{31}, S_{(0)}^{12}$, which will coincide with the components of the spin vector, and $S_{(0)}^{14} = S_{(0)}^{24} = S_{(0)}^{34} = 0$, but the same thing will not be true in the other Galilean systems. The tensor $S^{\alpha\beta}$ will then be the tensor \mathbf{S} that corresponds to the spin for the proper observer, as we have defined it before; it will not correspond to the spin in the other Galilean systems. That point seems to have escaped van Weysenhoff.

From the preceding definition, upon remarking that:

$$d_{i_0} G^\alpha = \int (D_{i_0} \mathcal{G}^\alpha) d\tau_0,$$

and upon taking into account the relations $D_{i_0} \mathcal{G}^\alpha = 0$ and $s^{\alpha\beta} u_\beta = 0$, one will infer that:

$$\begin{aligned} \dot{G}^\alpha &= 0, & S^{\alpha\beta} u_\beta &= 0, & \dot{S}^{\alpha\beta} &= G^\alpha u^\beta - G^\beta u^\alpha, \\ G^\alpha &= m_0 u^\alpha + \frac{1}{c^2} S^{\alpha\beta} \dot{u}_\beta, & m_0 &= -\frac{1}{c^2} u_\beta G^\beta. \end{aligned}$$

One then deduces that:

$$\begin{aligned} m_0 \dot{u}^\alpha + \frac{1}{c^2} S^{\alpha\beta} \ddot{u}_\beta &= 0, \\ \dot{S}^{\alpha\beta} &= \frac{1}{c^2} S^{\alpha\sigma} \dot{u}_\sigma u^\beta - \frac{1}{c^2} S^{\beta\sigma} \dot{u}_\sigma u^\alpha. \end{aligned}$$

In the first equation, one has suppressed a term in $\dot{S}^{\alpha\beta} \dot{u}_\beta$; indeed, it is zero, because one has:

$$\dot{S}^{\alpha\beta} \dot{u}_\beta = G^\alpha u^\beta \dot{u}_\beta - G^\beta u^\alpha \dot{u}_\beta,$$

but

$$u^\beta \dot{u}_\beta = 0,$$

from the fact that $u^\beta u_\beta = -c^2$, and:

$$G^\beta \dot{u}_\beta = m_0 \dot{u}_\beta u^\beta + \frac{1}{c^2} S^{\alpha\beta} \dot{u}_\alpha \dot{u}_\beta,$$

which is zero by reason of the antisymmetry of $S^{\alpha\beta}$:

From the fact that $m_0 = -1/c^2 u_\beta G^\beta$, one will infer that $\dot{m}_0 = -\frac{1}{c^2} u_\beta \dot{G}^\beta - \frac{1}{c^2} \dot{u}_\beta G^\beta$, and that $\dot{u}_\beta G^\beta$ is zero, as well as \dot{G}^β , in such a way that $\dot{m}_0 = 0$. Hence, m_0 is a constant, namely, the rest mass of the particle.

One further finds that $\dot{S}^{\alpha\beta} S_{\alpha\beta} = 0$, since $S^{\alpha\beta} u_\beta = 0$, so:

$$S^{\alpha\beta} S_{\alpha\beta} = \mathbf{S} \cdot \mathbf{S} - \mathbf{Q} \cdot \mathbf{Q} = S_0^2 = \text{const.}$$

The proper angular momentum of the particle is constant.

I shall not present the part of the Weyssenhoff article that was dedicated to the integration of the equations. That integration will lead one to consider the spinning particle as being animated in its proper system with a circular motion that is perpendicular to the vector $\mathbf{S}^{(0)}$. That image is certainly interesting, and can be useful in making a comparison with Schrödinger's *Zitterbewegung*. However, as the author himself has confirmed, it presents some difficulties when one compares it to Dirac's theory, and it does not present the quantum character that is currently indispensable for the representation of the properties of spinning corpuscles. By its intuitive character, it can even suggest false ideas in that subject.

We shall first introduce the electromagnetic tensor F such that:

$$\mathbf{H} = (F^{23}, F^{31}, F^{12}) \quad \text{and} \quad \mathbf{E} = (F^{14}, F^{24}, F^{34}),$$

and the tensor $m^{\alpha\beta}$ of *density of proper electric and magnetic moment*, which, as we know, is proportional to $s^{\alpha\beta}$ in any reference system:

$$m^{\alpha\beta} = \chi s^{\alpha\beta},$$

with $\chi = -e / m_0 c$ for the electron.

Weyssenhoff proposed that the energy of the particle in an electromagnetic field should be:

$$U = -\mathfrak{M} \cdot \mathbf{H} - \mathfrak{P} \cdot \mathbf{E} = -\frac{1}{2} F^{\alpha\beta} m_{\alpha\beta}.$$

The last expression seems inexact to us, since $m^{\alpha\beta}$ is the *density* of electromagnetic moment, and not that moment itself. One agrees to set:

$$U = -\frac{1}{2} \int F^{\alpha\beta} m_{\alpha\beta} d\tau.$$

An interesting situation intervenes here: $U = -\mathfrak{M} \cdot \mathbf{H} - \mathfrak{P} \cdot \mathbf{E}$ has the physical dimensions and meaning of energy. Meanwhile, it does not have the relativistic variance of energy, because it is not the temporal component of a quadri-vector. If one assumes the Weyssenhoff expression:

$$(IV.g) \quad U = -\frac{1}{2} F^{\alpha\beta} m_{\alpha\beta}$$

then U will be an invariant. However, upon taking the correct expression:

$$U = -\frac{1}{2} \int F^{\alpha\beta} m_{\alpha\beta} d\tau = -\frac{1}{2} \int F^{\alpha\beta} m_{\alpha\beta} d\tau_0 \sqrt{1 - \beta^2},$$

one will see that U has the form:

$$U = U_0 \sqrt{1 - \beta^2},$$

in which U_0 is the invariant $-\frac{1}{2} \int F^{\alpha\beta} m_{\alpha\beta} d\tau_0$. U transforms like a volume. If the particle is small enough for it to be assimilated to a point then one can set:

$$\dot{G}_\alpha = \frac{\mathcal{E}}{c} F_{\alpha\beta} u^\beta + \partial_\alpha \left(\frac{1}{2} \int F^{\rho\sigma} m_{\rho\sigma} d\tau_0 \right),$$

in which $\mathcal{E}/c F_{\alpha\beta} u^\beta$ is the (four-dimensional) Lorentz force, and ∂_α indicates the derivation with respect to the coordinates of the particle, when considered as a unit (e.g., the coordinates of the center of gravity of the particle). The first three equations give (since $dt_0 = dt \sqrt{1 - \beta^2}$):

$$\frac{dG_i}{dt} = f_i - \frac{\partial}{\partial x_0} \left(U_0 \sqrt{1 - \beta^2} \right) = f_i - \partial_i U, \quad i = 1, 2, 3,$$

(f = four-dimensional Lorentz force), and the fourth one expresses the conservation of energy:

$$\frac{dW}{dt} = (f_i \cdot v_i) - \frac{\partial}{\partial t} U.$$

In order to express the couple that is exerted upon the particle, one writes:

$$\dot{S}_{\alpha\beta} = G_\alpha u_\beta - G_\beta u_\alpha + \int (m_{\alpha\sigma} F_\beta^\sigma - m_{\beta\sigma} F_\alpha^\sigma) d\tau_0.$$

We have added an integral that does not appear in Weysenhoff's work.

Since:

$$m_0 = \int \mu_0 d\tau_0 = -\frac{1}{c^2} u^\beta \int \mathcal{G}_\beta d\tau_0,$$

one has:

$$\dot{m}_0 = -\frac{1}{c^2} u^\beta \int \dot{\mathcal{G}}_\beta d\tau_0 = -\frac{1}{c^2} \dot{G}^\beta u_\beta = -\frac{1}{c^2} \dot{G}_\beta u^\beta = d_{t_0} \left(-\frac{1}{2c^2} \int m^{\rho\sigma} F_{\rho\sigma} d\tau_0 \right).$$

The quantity:

$$\begin{aligned} m_{00} &= m_0 + \frac{1}{2c^2} \int F_{\rho\sigma} m^{\rho\sigma} d\tau_0 = -\frac{1}{c^2} G^\beta u_\beta + \frac{1}{2c^2} \int F_{\rho\sigma} m^{\rho\sigma} d\tau_0 \\ &= m_0 - \frac{U_0}{c^2} = m_0 - \frac{U}{c\sqrt{1 - \beta^2}}, \end{aligned}$$

which reduces to m_0 when the fields are zero, is a constant of the motion. The proper mass m_0 is the sum of a constant part m_{00} and the variable part U_0 / c^2 . Upon letting m'_0

denote what Weyssenhoff called m_0 , and letting m_0 denote what he called m_{00} , we will then have:

$$m'_0 = m_0 + \frac{U}{c^2},$$

which defines a variable proper mass m'_0 of the spinning particle in the electromagnetic field that we ultimately recover.

Remark. – The relation:

$$(IV.f) \quad S_{\alpha\beta} u^\beta = 0,$$

when combined with the definition:

$$(IV.g) \quad \mu^{\alpha\beta} = \chi s^{\alpha\beta},$$

will give us:

$$\mu_{\alpha\beta} u^\beta = 0.$$

Upon denoting the quadri-vector $\pi^{(0)} = m_0 c \mathbf{u}$ by $\pi^{(0) (1)}$, we will get:

$$(IV.h) \quad \mu_{\alpha\beta} \pi^{(0)\beta} = 0,$$

which is a relation that we shall ultimately utilize.

(¹) Which is not to be confused with the density of proper electric moment.

CHAPTER V

DIRAC'S THEORY OF THE SPINNING ELECTRON

1. THE WAVE EQUATIONS FOR THE SPINNING ELECTRON

Dirac found the fundamental equations for his theory by seeking to construct a wave mechanics of the electron that was relativistic and would permit one to preserve a form for the probability of presence that would be analogous to the positive-definite $|\psi|^2$ that was valid in the original non-relativistic wave mechanics.

Inspired by a previous attempt by Pauli, he assumed that the wave function ψ of the electron must have several components ψ_k , and that the probability density of presence must be positive-definite, which is expressed as a function of the ψ_k by:

$$\rho = \sum_k |\psi_k|^2.$$

In order for the probability of all possible positions of the electron to be equal to 1, one must then normalize the function by setting:

$$\int_D \sum_k |\psi_k|^2 d\tau = 1,$$

in which D is the spatial domain in which one finds the electron. However, that condition will be acceptable only if it remains true for all time as a result of the wave equations once it is true at a given instant. Now, Dirac pointed out that in order for that to be true, the equations that are satisfied by the ψ_k must be of first order in t , since it is necessary that being given only the ψ_k at an initial instant must suffice in order to determine all of their later evolution. The relativistic symmetry between time and spatial coordinates then indicates that one must look for a system of partial differential equations in the ψ_k that has first order in the space and time variables.

Upon arriving at that conclusion, Dirac showed that one must take at least four functions ψ_k , and he assumed that one must limit oneself to four of them. Now, at the beginning of the development of wave mechanics, several authors (viz., de Donder, Fock, Gordon, and L. de Broglie) simultaneously gave a relativistic form to the wave equation. In order to do that, they started by remarking that in the old relativistic dynamics, in the absence of a field, the energy W of a corpuscle is linked with the components of its impulse \mathbf{p} and its proper mass m_0 by the formula:

$$\frac{W^2}{c^2} - (p_x^2 + p_y^2 + p_z^2 + m_0^2 c^2) = 0.$$

Replace W in that equation with the operator $\frac{h}{2\pi i} \frac{\partial}{\partial t}$, and each p_i with the operator $\frac{-h}{2\pi i} \frac{\partial}{\partial q_i}$. One will then get an operator, and upon applying that operator to the function ψ and equating the result to zero, one will get:

$$\square\psi + \frac{4\pi^2}{h^2} m_0^2 c^2 \psi = 0, \quad \text{with} \quad \square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta,$$

which is an equation that was proposed as the wave equation of wave mechanics with just one wave function in the absence of a field.

That attempt did not give satisfactory results (notably, for the calculation of the hydrogen levels), and Dirac's arguments show that such a second-order equation cannot serve as a satisfactory basis for the wave mechanics of the electron, and that one must write four first-order equations in ψ_k . Nevertheless, it is natural to think that each of the four ψ_k must satisfy the preceding second-order equation, at least in the absence of an external field. Briefly, Dirac was led to seek four equations of first-order in x, y, z, t for the four ψ_k that would be valid in the absence of a field and would imply the equation:

$$\square\psi_k + \frac{4\pi^2}{h^2} m_0^2 c^2 \psi_k = 0$$

for each of the ψ_k .

In order to write those first-order equations, one will be led to employ four Hermitian matrices $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ with four rows and four columns, and one will define the symbol $\alpha_i \psi_k$ by:

$$\alpha_i \psi_k = \sum_{j=1}^4 (\alpha_i)_{kj} \psi_j,$$

in such a way that applying the operator α_i to the ψ_k amounts to forming a certain linear combination of the four ψ_k .

One imposes the following conditions upon the four matrices α_i :

1. The square of each of them must be equal to the identity matrix.
2. Two different matrices will anti-commute.

Those two conditions can be summarized in the single formula:

$$(V.a) \quad \alpha_i \alpha_j + \alpha_j \alpha_i = 2 \delta_{ij} \cdot 1,$$

in which 1 represents the identity matrix with four rows and four columns:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Furthermore, those conditions are not sufficient to determine the matrices α_i completely, but we shall see that the indeterminacy that persists in the matrices α_i does not imply any indeterminacy in the physical consequences of the theory.

The Dirac equations can be written symbolically with the aid of the α_i matrices as:

$$(V.b) \quad \left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\frac{\partial}{\partial x} \alpha_1 + \frac{\partial}{\partial y} \alpha_2 + \frac{\partial}{\partial z} \alpha_3 + \kappa m_0 c \alpha_4 \right) \right] \psi_k = 0,$$

with $k = 1, 2, 3, 4$, and $\kappa = 2\pi i / h$. They represent four simultaneous equations in the ψ_k .

If one applies the operator:

$$\left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\frac{\partial}{\partial x} \alpha_1 + \frac{\partial}{\partial y} \alpha_2 + \frac{\partial}{\partial z} \alpha_3 + \kappa m_0 c \alpha_4 \right) \right]$$

to the left-hand side of (V.b) then one will verify that upon taking (V.a) into account, one will find that for each of the ψ_k :

$$\square \psi_k + \frac{4\pi^2}{h^2} m_0^2 c^2 \psi_k = 0, \quad k = 1, 2, 3, 4.$$

It remains for us to explain why the fact that the matrices α_i remain arbitrary to a large extent does not imply any indeterminacy in the physical consequences of (V.b). In order to show that, we will start from the following result: One can show that if $\alpha_1, \dots, \alpha_4$, and $\alpha'_1, \dots, \alpha'_4$ are two sets of Hermitian matrices that each satisfy the equations (V.a) then it will be possible to find a *unitary* matrix S with four rows and four columns such that one has:

$$\alpha'_i = S^{-1} \alpha_i S = S^\dagger \alpha_i S.$$

In other words, one can always pass from the α_i to the α'_i by a canonical transformation that preserves the Hermitian character and the commutation relations (V.a).

From the Dirac equations, when they are written in terms of the α'_i :

$$\left[\frac{1}{c} \frac{\partial}{\partial t} - \left(\frac{\partial}{\partial x} \alpha'_1 + \frac{\partial}{\partial y} \alpha'_2 + \frac{\partial}{\partial z} \alpha'_3 + \kappa m_0 c \alpha'_4 \right) \right] \psi'_k = 0,$$

or furthermore:

$$0 = \left[\frac{1}{c} \frac{\partial}{\partial t} - \left(\frac{\partial}{\partial x} S^{-1} \alpha_1 S + \frac{\partial}{\partial y} S^{-1} \alpha_2 S + \frac{\partial}{\partial z} S^{-1} \alpha_3 S + \kappa m_0 c S^{-1} \alpha_4 S \right) \right] \psi'_k,$$

one will infer immediately, upon multiplying the latter on the left by S , that the quantities $S\psi'_k$ are solutions of the Dirac equations, when they are written in terms of the α_i . The quantities:

$$\psi_k = S\psi'_k \quad (k = 1, 2, 3, 4)$$

are linear combination of the ψ'_k . Hence, changing the α_i amounts to making a unitary transformation of the ψ'_k . Now, as we will see, the physical quantities that one encounters in Dirac's theory all have the form $\sum_{k=1}^4 \psi_k^* A \psi_k$ or $\int_D \sum_{k=1}^4 \psi_k^* A \psi_k d\tau$, in which A is a linear, Hermitian operator that can contain the α_i or their products, and in turn, operate upon the indices k of the ψ_k . When one passes from the representation α_i to a system α'_i (one then says that one passes from the representation α_i to the representation α'_i), the operator A will become $A' = S^{-1} A S$, and upon taking into account that $S^{-1} = S^\dagger$, the quantity $\sum_{k=1}^4 \psi_k^* A \psi_k$ will become:

$$\sum_{k=1}^4 \psi_k^* A \psi_k = \sum_{k=1}^4 S^* \psi_k'^* A S \psi'_k = \sum_{k=1}^4 \psi_k'^* S^{-1} A S \psi'_k = \sum_{k=1}^4 \psi_k'^* A' \psi'_k.$$

Hence, the quantities that have a physical sense will keep the same value when one changes the representation. Thanks to that fact, the partial indeterminacy in the α_i will not imply any indeterminacy in the physical predictions that are confirmed for Dirac's theory, which is obviously necessary for that indeterminacy in the α_i to be acceptable.

The Dirac equations can be written in the condensed form:

$$\frac{1}{\kappa} \frac{\partial \psi}{\partial t} = \frac{h}{2\pi i} \frac{\partial \psi}{\partial t} = \mathcal{H} \psi,$$

in which H is the Dirac Hamiltonian operator, which is given by:

$$\mathcal{H} = \frac{c}{\kappa} \left(\frac{\partial}{\partial x} \alpha_1 + \frac{\partial}{\partial y} \alpha_2 + \frac{\partial}{\partial z} \alpha_3 + \kappa m_0 c \alpha_4 \right).$$

In the general case where the electron moves in an electromagnetic field that is derived from a scalar potential V and the potential vector \mathbf{A} , Dirac replaced the equations of propagation that are valid in the absence of a field with the following ones:

$$\left(\frac{1}{c}\frac{\partial\psi}{\partial t} + \kappa\frac{e}{c}V\right)\psi_k = \left[\left(\frac{\partial}{\partial x} + \kappa\frac{e}{c}A_x\right)\alpha_1 + \left(\frac{\partial}{\partial y} + \kappa\frac{e}{c}A_y\right)\alpha_2 + \left(\frac{\partial}{\partial z} + \kappa\frac{e}{c}A_z\right)\alpha_3 + \kappa m_0 c \alpha_4\right]\psi_k,$$

for $k = 1, 2, 3, 4$. (The charge of the electron is denoted by $-e$.)

Profiting from the indeterminacy in the α_i , we shall generally make use of the following matrices, whose use is convenient:

$$\alpha_1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}, \quad \alpha_2 = \begin{vmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{vmatrix}, \quad \alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix}, \quad \alpha_4 = \begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix},$$

whose Hermitian character is obvious, and which verify the relations (V.a).

With that choice of the α_i , the Dirac equations can be written:

$$(V.c) \quad \begin{cases} \frac{1}{c}\frac{\partial\psi_1}{\partial t} = \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\psi_4 + \frac{\partial}{\partial z}\psi_3 - \frac{2\pi i}{h}m_0c\psi_1, \\ \frac{1}{c}\frac{\partial\psi_2}{\partial t} = \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\psi_3 - \frac{\partial}{\partial z}\psi_4 - \frac{2\pi i}{h}m_0c\psi_2, \\ \frac{1}{c}\frac{\partial\psi_3}{\partial t} = \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\psi_2 + \frac{\partial}{\partial z}\psi_1 + \frac{2\pi i}{h}m_0c\psi_3, \\ \frac{1}{c}\frac{\partial\psi_4}{\partial t} = \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\psi_1 - \frac{\partial}{\partial z}\psi_2 + \frac{2\pi i}{h}m_0c\psi_4. \end{cases}$$

One might be surprised at the highly asymmetric form of these equations, in which the z -axis plays a special role. In order to understand the meaning of that asymmetry, one must recall that the role of the wave functions is essentially to permit one to evaluate certain probabilities. Now, for the spinning electron, questions of probability must be posed with respect to a certain reference axis D : For example, one can demand to know what the possible values are for the components of the spin in the reference direction D and the probabilities of those possible values. The choice of α_i that we have made corresponds to the case in which the reference direction D coincides with the z -axis. If one would like to address a question of the probability of the value of spin that is posed for a reference direction that does not coincide with the initially-chosen direction for the z -axis then one must first make a change of axis that takes Oz to the direction D , and it will be the new ψ'_k (which, as we shall see later, are linear combinations of the old ψ_k) that provide the desired probabilities.

2. THE RELATIVISTIC INVARIANCE OF THE DIRAC EQUATIONS

Dirac showed that if one makes a change of Galilean reference system by subjecting the coordinates to a Lorentz transformation then the equations of propagation will keep the same form in the new system that they had in the old one, while the components ψ_k of the wave function will submit to a linear transformation of the form:

$$\psi_k = \sum_{l=1}^4 \Lambda_{kl} \psi_l, \quad k = 1, 2, 3, 4.$$

The coefficients Λ_{ij} are the elements of a matrix Λ with four rows and four columns that is neither Hermitian nor unitary in the general case. Naturally, the matrix Λ depends upon the Lorentz transformation that was performed, but one must note that the transformation of the four ψ_k is not the one that the four components of a space-time quadri-vector are subjected to. Although the mathematical entity ψ of Dirac's theory has four components, it is not a quadri-vector. It belongs to a category of mathematical entities that had not been introduced into physics before Dirac, and that one now studies in the theory of spinors. In order to find quantities in Dirac's relativistic wave mechanics that will transform like the components of a tensor under a Lorentz transformation, one must form certain bilinear combination of the ψ_k and the ψ_k^* that we shall speak of at length later on.

In order to prove the invariance of the Dirac equations, we shall appeal to the world variables $x_1 = x, x_2 = y, x_3 = z, x_4 = ict$, and we set, by definition:

$$P_j = \frac{-h}{2\pi i} \frac{\partial}{\partial x_j} - \frac{e}{c} A_j \quad (j = 1, 2, 3), \quad P_4 = \frac{-h}{2\pi i} \frac{\partial}{\partial x_4} + \frac{e}{c} iV.$$

We then write the Dirac equations in the symbolic form:

$$\left(\frac{P_4}{i} + \sum_{j=1}^3 P_j \alpha_j + m_0 c \alpha_4 \right) \psi_k = 0,$$

so after multiplying them by $i \alpha_4$ on the left:

$$\left(\alpha_4 P_4 + \sum_{j=1}^3 i \alpha_4 \alpha_j P_j + i m_0 c \right) \psi_k = 0.$$

With von Neumann, we now replace the matrices α_i with the matrices γ_i that are defined by the equations:

$$\gamma_1 = i \alpha_4 \alpha_1, \quad \gamma_2 = i \alpha_4 \alpha_2, \quad \gamma_3 = i \alpha_4 \alpha_3, \quad \gamma_4 = \alpha_4.$$

It is easy to see that one also has:

$$\gamma_i \gamma_j + \gamma_j \gamma_i = 2 \delta_{ij} \cdot 1,$$

just as one had for the α_i . The Dirac equations can then be written symbolically as:

$$\left(\sum_{j=1}^4 P_j \gamma_j + i m_0 c \right) \psi_k = 0.$$

That elegant form agrees with the relativistic idea that the spatial and temporal coordinates must always play a symmetric role. By contrast, it has the inconvenience that it exhibits the special role that time plays in quantum theories in the equations in α_i .

Now suppose that we change Galilean axes, so the space-time coordinates will be subjected to a Lorentz transformation. It is well-known in relativity that such a transformation is equivalent to a rotation of the axes in Minkowski space. The new variables x'_4 after the transformation will then be linked with the old ones by the formulas:

$$x_i = \sum_j o_{ij} x'_j,$$

in which o is a matrix with four rows and four columns. o is not a real matrix, due to the pure imaginary character of the world-variable x_4 ; those of its elements that contain the index 4 once will be pure imaginary. Moreover, one has the orthogonality relation:

$$\sum_l o_{li} o_{lj} = \sum_l o_{il} o_{jl} = \delta_{ij}.$$

It is clear that the P_i transform like the x_i ; i.e.:

$$P_i = \sum_j o_{ij} P'_j.$$

After the change of axes, the Dirac equations will then be written:

$$\left(\sum_{i=1}^4 \sum_{j=1}^4 o_{ij} P'_j + i m_0 c \right) \psi_k = 0,$$

in which the components ψ_k and ψ can be expressed with the aid of the new variables x'_i . If we set:

$$(V.d) \quad \gamma'_j = \sum_{i=1}^4 o_{ij} \gamma_i$$

then we can write:

$$\left(\sum_{j=1}^4 \gamma'_j P'_j + i m_0 c \right) \psi = 0.$$

Equation (V.d) expresses the matrices γ'_j as functions of the matrices γ_i . If one would like to make its meaning more precise then one can write:

$$(\gamma'_j)_{mn} = \sum_{i=1}^4 o_{ij} (\gamma_i)_{mn},$$

in which $(\gamma_i)_{mn}$, for example, is the mn element of γ_i .

Since one has:

$$(\gamma'_j)_{mn}^* = \sum_{i=1}^4 o_{ij}^* (\gamma_i)_{mn}^* = \sum_{i=1}^4 o_{ij}^* (\gamma_i)_{mn} \neq (\gamma'_j)_{mn},$$

one sees that the γ'_j are not Hermitian, because the o_{ij} are not all real (i.e., $o_{ij}^* \neq o_{ij}$).

By contrast, it is easy to verify that:

$$\gamma'_i \gamma'_j + \gamma'_j \gamma'_i = \sum_{k,l} o_{ki} o_{lj} (\gamma_k \gamma_l + \gamma_l \gamma_k) = \sum_{k,l} o_{ki} o_{lj} 2\delta_{kl} 1 = 2 \sum_k o_{ki} o_{kj} 1 = 2 \delta_{ij} \cdot 1.$$

If the γ'_k are Hermitian then, from a general result that was stated before, they will be linked with the γ_i by the formula:

$$\gamma'_i = S^{-1} \gamma_i S,$$

in which S is unitary. However, that cannot be true, since canonical transformations will preserve the Hermitian character, and the γ'_j are not Hermitian. By contrast, we can have a relation of the form:

$$(V.e) \quad \gamma'_i = \Lambda^{-1} \gamma_i \Lambda,$$

in which Λ is a matrix that is not unitary, in general (i.e., $\Lambda^{-1} \neq \Lambda^+$). That transformation of the γ obviously preserves the commutation relations, but it does not preserve the Hermitian character.

We shall, for the moment, assume the validity of the relations (V.e), and reserve its proof until later. In coordinates x'_j , the Dirac equation can then be written:

$$\left(\sum_{j=1}^4 \Lambda^{-1} \gamma_j \Lambda P'_j + im_0 c \right) \psi = 0.$$

Multiply this on the left by Λ and note that the matrix Λ , which corresponds to an operation that is performed on the indices of the ψ_k , commutes with P'_j . We will have:

$$\left(\sum_{j=1}^4 \gamma_j P'_j + im_0 c \right) \Lambda \psi = 0.$$

It then results that the function $\psi' = \Lambda\psi$, whose components are $\psi'_k = \sum_j \Lambda_{kj} \psi_j$, is a solution of the Dirac equations in the system of primed variables with the same matrices γ_i (or α_i) that it had in the original system. That is in fact the result that we would like to establish.

However, it will remain for us prove the relation (V.e). In order to do that, we first remark that those transformations form a group, because if one has:

$$\gamma'_i = \Lambda_1^{-1} \gamma_i \Lambda_1, \quad \gamma''_i = \Lambda_2^{-1} \gamma'_i \Lambda_2$$

then one will also have:

$$\gamma''_i = \Lambda_2^{-1} \Lambda_1^{-1} \gamma_i \Lambda_1 \Lambda_2 = (\Lambda_1 \Lambda_2)^{-1} \gamma_i (\Lambda_1 \Lambda_2).$$

One easily concludes that if the formula is true for an infinitesimal rotation of the world-axes then it will also be true for a finite rotation; it will suffice to prove it for an infinitesimal rotation. Now, for such a rotation, one can set:

$$o_{ij} = \delta_{ij} + \varepsilon_{ij},$$

in which the ε_{ij} are very small quantities whose squares can be neglected. In order for the orthogonality condition to be satisfied, it is necessary that:

$$\varepsilon_{ij} = -\varepsilon_{ji}, \quad \varepsilon_{ij} = 0.$$

The matrix Λ differs slightly from the matrix 1, and we can set:

$$\Lambda = 1 + \frac{1}{2} \sum_{k,l} \varepsilon_{kl} T_{kl}, \quad T_{kl} = -T_{lk},$$

in which the ε are very small, so, to second order in ε :

$$\Lambda^{-1} = 1 - \frac{1}{2} \sum_{k,l} \varepsilon_{kl} T_{kl}.$$

The T_{kl} are matrices whose elements are unknown. We must prove that one can choose the T_{kl} in such a fashion that we will have:

$$\sum_j o_{ji} \gamma_j = \Lambda^{-1} \gamma_i \Lambda;$$

i.e.:

$$\begin{aligned} \gamma_i + \sum_j \varepsilon_{ji} \gamma_j &= \left(1 - \frac{1}{2} \sum_{k,l} \varepsilon_{kl} T_{kl} \right) \gamma_i \left(1 + \frac{1}{2} \sum_{k,l} \varepsilon_{kl} T_{kl} \right) \\ &= \gamma_i + \frac{1}{2} \sum_{k,l} \varepsilon_{kl} (\gamma_i T_{kl} - T_{kl} \gamma_i) + \dots, \end{aligned}$$

up to terms in ε^2 , namely:

$$\sum_j \varepsilon_{ji} \gamma_j = \frac{1}{2} \sum_{k,l} \varepsilon_{kl} (\gamma_i T_{kl} - T_{kl} \gamma_i) \quad (i = 1, 2, 3, 4).$$

That matrix equation admits the solution:

$$T_{kl} = -T_{lk} = -\frac{1}{2} \gamma_k \gamma_l.$$

We have thus defined the existence of the matrix Λ for an infinitesimal rotation of the world-axes, and we have found its expression. As we have said, the existence of a general matrix Λ for an arbitrary Lorentz transformation will then follow, and the proof of the relativistic invariance of the Dirac equations is thus found to have been achieved.

The matrix Λ is unitary only if the transformation of the axes that is imagined reduces to a simple rotation of the spatial axes with no relative motion. For a change of axes with relative motion, Λ will not be unitary (indeed, one can show that $\Lambda = \gamma_4 \Lambda^{-1} \gamma_4$, although $\Lambda^+ = \Lambda^{-1}$). It will then result that:

$$\sum_k \psi_k^* \psi'_k = \sum_k (\Lambda \psi_k)^* \Lambda \psi_k = \sum_k \psi_k^* \Lambda^+ \Lambda \psi_k$$

is not equal to $\sum_k \psi_k^* \psi_k$, because $\Lambda^+ \Lambda = \gamma_4 \Lambda^{-1} \gamma_4 \Lambda \neq 1$. The probability of presence is not invariant for a Lorentz transformation with relative motion then; indeed, we see that it is the temporal component of a quadri-vector.

3. THE SPIN OF THE ELECTRON IN DIRAC’S THEORY

What is remarkable in Dirac’s theory is that when one starts from the attempt to constitute a relativistic wave mechanics of the electron, without addressing the question of spin, one finds that one has *automatically* introduced the existence of the proper kinetic moment and proper magnetic moment that are required for the interpretation of the experimental facts.

In non-relativistic wave mechanics, we found the following result, which is completely analogous to a result in classical mechanics:

In a central force field, the quantities that are the “components of the orbital kinetic moment M ,” which correspond to the operators $(x p_y - y p_x)$, ..., are first integrals.

On the contrary, in the wave mechanics of the Dirac electron, the components of the orbital kinetic moment do not commute with the Hamiltonian, and are *not* first integrals. In order to obtain first integrals in Dirac’s theory – i.e., operators that commute with the Hamiltonian H – one will be led to add some new quantities to each of the components of the orbital moment, namely, the components of the proper kinetic moment, or “spin” of the electron, which are components that are defined in the quantum manner by the operators:

$$(S_x)_{\text{op}} = \frac{h}{4\pi} i \alpha_2 \alpha_3, \quad (S_y)_{\text{op}} = \frac{h}{4\pi} i \alpha_3 \alpha_1, \quad (S_z)_{\text{op}} = \frac{h}{4\pi} i \alpha_1 \alpha_2,$$

whose proper values are, one easily verifies, $\pm h / 4\pi$. It is the vector $\mathbf{M} + \mathbf{S}$ that is a first integral in a central field here, and which will then play the role of *total* kinetic moment of the electron: It is the sum of the orbital kinetic moment and the spin. Each of the components of the spin \mathbf{S} do, in fact, have the possible values $\pm h / 4\pi$ that are predicted by the Uhlenbeck-Goudsmit hypothesis.

If we write the operator \mathbf{S}_{op} in the form that was employed before:

$$\mathbf{S}_{\text{op}} = \frac{h}{2\pi} \mathbf{s}_{\text{op}}$$

then we will see that we must set:

$$(s_x)_{\text{op}} = i \alpha_2 \alpha_3, \quad (s_y)_{\text{op}} = i \alpha_3 \alpha_1, \quad (s_z)_{\text{op}} = i \alpha_1 \alpha_2$$

for the spin of the electron, and with those definitions, we will find the non-commutation relations:

$$[s_x, s_y] = -i s_z, \quad [s_y, s_z] = -i s_x, \quad [s_z, s_x] = -i s_y.$$

Those are, in fact, the relations that are required by the general quantum theory of kinetic moments, which shows the coherence of the adopted definitions.

The corresponding operator $s^2 = s_x^2 + s_y^2 + s_z^2$ is equal to:

$$s^2 = -\frac{1}{4} [(\alpha_1 \alpha_2)^2 + (\alpha_2 \alpha_3)^2 + (\alpha_3 \alpha_1)^2],$$

which corresponds to the value $s = 1/2$; i.e., to the proper value $s(s+1) = 3/4$ of s^2 .

We have seen that the proper kinetic moment of the electron is associated with a proper magnetic moment. Dirac's theory also arrives at that proper magnetic moment, and also in that form. Upon seeking the second-order equations that replace the equations:

$$\square \psi_k + \frac{4\pi^2}{h^2} m_0^2 c^2 \psi_k = 0$$

in the case for which there exists an external electromagnetic field, one will arrive at the conclusion that the electron behaves as if it were endowed with a proper magnetic moment \mathfrak{M} and a proper electric moment \mathfrak{P} . That conclusion is related to the Uhlenbeck-Goudsmit hypothesis, since under those hypotheses an electron that possesses a magnetic moment in its proper system will, by a relativistic effect, also possess a proper electric moment in a Galilean system in which it is motion. As always in wave mechanics, the components of \mathfrak{M} and \mathfrak{P} must be defined by operators. The argument that we shall develop leads us to set:

$$(\mathfrak{M}_x)_{\text{op}} = B i \alpha_2 \alpha_3 \alpha_4, \quad (\mathfrak{M}_y)_{\text{op}} = B i \alpha_3 \alpha_1 \alpha_4, \quad (\mathfrak{M}_z)_{\text{op}} = B i \alpha_1 \alpha_2 \alpha_4,$$

$$(\mathfrak{P}_x)_{\text{op}} = B i \alpha_1 \alpha_4, \quad (\mathfrak{P}_y)_{\text{op}} = B i \alpha_2 \alpha_4, \quad (\mathfrak{P}_z)_{\text{op}} = B i \alpha_3 \alpha_4,$$

in which B denotes the Bohr magneton $B = \frac{-eh}{4\pi m_0 c}$.

The six operators all have proper values of $\pm B$, and that conclusion will also agree with the Uhlenbeck-Goudsmit hypothesis about the value of the proper magnetic moment of the electron.

In order to justify the form of the operators that were indicated above, we set:

$$p_4 = \frac{h}{2\pi i} \cdot \frac{1}{c} \cdot \frac{\partial}{\partial t}, \quad p_j = -\frac{h}{2\pi i} \cdot \frac{\partial}{\partial x_j} \quad (j = 1, 2, 3),$$

and we first write the equation of non-relativistic wave mechanics for the electron in the form:

$$\left[cp_4 - \left(\frac{1}{2m_0} \sum_{j=1}^3 p_j^2 + U \right) \right] \psi = 0,$$

or further:

$$(V.f) \quad (2m_0 c p_4 - \sum_{j=1}^3 p_j^2 - 2m_0 U) \psi = 0.$$

On the other hand, we have seen that for a free electron, a theory of wave mechanics that employs just one function ψ must write:

$$(V.g) \quad (p_4^2 - \sum_{j=1}^3 p_j^2 - m_0^2 c^2) \psi = 0.$$

For the electron that is placed in an electromagnetic field, one must replace the operators p with the following ones:

$$P_4 = p_4 + \frac{e}{c} V, \quad P_j = p_j + \frac{e}{c} A_j \quad (j = 1, 2, 3),$$

and one will get the equation of propagation:

$$(P_4^2 - \sum_{j=1}^3 P_j^2 - m_0^2 c^2) \psi = 0,$$

which will give back equation (V.f) in the non-relativistic approximation.

In Dirac's theory, one starts with the system of equations:

$$(P_4 + \sum_{j=1}^3 \alpha_j P_j - \alpha_4 m_0 c) \psi_k = 0 \quad (k = 1, 2, 3, 4).$$

If we apply the operator:

$$[P_4 - \sum_{j=1}^3 \alpha_j P_j - \alpha_4 m_0 c]$$

to that equation then we will easily get:

$$\left[P_4^2 - \sum_{j=1}^3 P_j^2 - m_0^2 c^2 + \sum_{j=1}^3 \alpha_j (P_4 P_j - P_j P_4) - \sum_{i \neq j=1}^3 (\alpha_i \alpha_j P_i P_j - \alpha_j \alpha_i P_j P_i) \right] \psi_k = 0.$$

Upon recalling that the fields are deduced from potentials by the formulas:

$$\mathbf{H} = \text{rot } \mathbf{A}, \quad \mathbf{E} = -\text{grad } V - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t},$$

one will find that:

$$0 = \left[P_4^2 - \sum_{j=1}^3 P_j^2 - m_0^2 c^2 + \frac{e}{c} \frac{h}{2\pi i} (\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z) - \frac{e}{c} \frac{h}{2\pi i} (\alpha_1 E_x + \alpha_2 E_y + \alpha_3 E_z) \right] \psi_k.$$

If only the first three terms in that equation exist then one will revert to equation (V.g), when it is applied to each of the ψ_k , and thus to equation (V.f) in the non-relativistic approximation. The new element that is introduced by Dirac's theory here is the presence of the last two terms, which one must interpret. In order to do that, we compare the latter equation with the non-relativistic equation (V.f). That comparison shows that the two terms to be interpreted must be considered (at least, in the non-relativistic approximation) to be the product of a factor $-2m_0$ with potential energy of interaction between the electromagnetic field of the electron. However, if one would not like to resort to the non-relativistic approximation then one must replace the proper mass m_0 in that statement with the quantity $\alpha_4 m_0$ that figures in the Dirac equations. Upon writing that $-2m_0 \alpha_4 U$ is equal to the last two terms of the latter equation, one will then see that U is the sum of the following terms:

$$U_e = \frac{eh}{4\pi m_0 c} i (\alpha_1 \alpha_4 E_x + \alpha_2 \alpha_4 E_y + \alpha_3 \alpha_4 E_z),$$

$$U_g = \frac{eh}{4\pi m_0 c} i (\alpha_2 \alpha_3 \alpha_4 H_x + \alpha_3 \alpha_1 \alpha_4 H_y + \alpha_1 \alpha_2 \alpha_4 H_z).$$

Now, a body that is endowed with a magnetic moment \mathfrak{M} and an electric moment \mathfrak{P} that is placed in an electromagnetic field that is defined by the vectors \mathbf{E} and \mathbf{H} will possess a potential energy that is equal to:

$$U = U_e + U_m = -(\mathfrak{M} \cdot \mathbf{H}) - (\mathfrak{P} \cdot \mathbf{E})$$

$$= - (\mathfrak{M}_x H_x + \mathfrak{M}_y H_y + \mathfrak{M}_z H_z) - (\mathfrak{P}_x E_x + \mathfrak{P}_y E_y + \mathfrak{P}_z E_z) .$$

Upon identifying the preceding expressions of U_e and U_m , one will recover the proper expressions for the operators that correspond to \mathfrak{M} and \mathfrak{P} . Hence, the magnetic moment and the electric moment seem to be contained in the Dirac equations themselves.

CHAPTER VI

FORMALISM AND PHYSICAL INTERPRETATION OF DIRAC'S THEORY

1. GENERAL FORMALISM OF RELATIVISTIC WAVE MECHANICS FOR THE DIRAC ELECTRON

In Dirac's wave mechanics, one recovers a formalism that is analogous to the one that we encountered in non-relativistic mechanics. All of the statements that were valid in the latter can be transposed into the former, under the condition that one must nonetheless assume that, along with the operators that act upon the coordinates, some other operators act upon the indices k in the functions ψ_k (such as the α_i matrices or linear, Hermitian combinations of them under addition and multiplication), and also under the condition that one must always include a summation over the index k from 1 to 4 in all of the formulas.

That is how the expressions that were given already for the probability density and the normalization formula for ψ are obtained by starting with the corresponding formulas of non-relativistic wave mechanics and adding a sum over k . Meanwhile, the definition of the *flux of the probability of presence* vector that one can combine with $\rho = \sum_{k=1}^4 |\psi_k|^2$ in order to define a *density-flux* quadri-vector has a particular form in Dirac's theory, because one can set:

$$\mathbf{f} = -c \sum_{k=1}^4 \psi_k^* \boldsymbol{\alpha} \psi_k .$$

$\boldsymbol{\alpha}$ denotes the matrix-vector whose three components are $\alpha_1, \alpha_2, \alpha_3$. In Dirac's theory, one easily shows that ρ and \mathbf{f} obey the conservation relation $\partial \rho / \partial t + \text{div } \mathbf{f} = 0$ by combining the equations of propagation and their conjugates. It will then result that $\int_D \rho d\tau$ is constant, by reason of the conditions that were imposed upon ψ already on the boundary of the domain D . Hence, if that integral is equal to 1 at an arbitrary instant then it will remain equal to 1, which will permit one to normalize ψ by setting:

$$\int_D \rho d\tau = \int_D \sum_k |\psi_k|^2 d\tau = 1.$$

In Dirac's wave mechanics, as in non-relativistic wave mechanics, any measurable quantity (i.e., observable) that is attached to a corpuscle will correspond to a linear, Hermitian operator A . In general, A can act upon not only the coordinates, but also upon the components indices of ψ_k . The operators will remain the same as before for the coordinates and components of the impulse, and they will act upon only the coordinates. The operator for energy will be the operator H that was defined above, which acts upon both the indices k and the variables x, y, z . For the new variables that are introduced in

Dirac's theory (e.g., proper kinetic moments, proper electric and magnetic moments), the operators will act upon only the index.

The possible values of the measurable quantity that corresponds to a linear, Hermitian operator A here are the proper values of the equation:

$$A \varphi_k = \alpha \varphi_k \quad k = 1, 2, 3, 4.$$

The proper value α_l corresponds to a proper function $\psi^{(l)}$ that has four components $\psi_1^{(l)}, \psi_2^{(l)}, \psi_3^{(l)}, \psi_4^{(l)}$. Those proper functions are orthogonal to each other (at least, if one chooses them conveniently in the case of degeneracy); i.e., one has:

$$\int_D \sum_{k=1}^4 \varphi_k^{(l)*} \varphi_k^{(l')} d\tau = 0 \quad (l \neq l').$$

Moreover, they form a complete system; i.e., one can always, for example, develop the ψ_k of the wave function ψ in the form:

$$\psi_k = \sum_l c_l \varphi_k^{(l)}, \quad k = 1, 2, 3, 4,$$

in which the c_l are independent of k . One often writes the four preceding equations in the condensed form:

$$\psi = \sum_l c_l \varphi^{(l)}.$$

The probability of the proper value α_l is given by $|c_l|^2$. One easily deduces from this that the mean value of A must be defined by:

$$\bar{A} = \sum_l c_l \varphi_k^{(l)} a_l = \int_D \sum_{k=1}^4 \psi_k^* A \psi_k d\tau.$$

Similarly, if $\psi_{i,1}, \dots, \psi_{i,4}$ are the components of the proper function of the Hamiltonian operator then the elements of the Heisenberg matrix that is generated by the operator A in the system of proper functions of H will be, by definition:

$$a_{ij} = \int_D \sum_{k=1}^4 \psi_{i,k}^* A \psi_{j,k} d\tau.$$

The mean-value densities are defined by $\sum_{k=1}^4 \psi_k^* A \psi_k$ and the densities of the matrix

elements $\sum_{k=1}^4 \psi_{i,k}^* A \psi_{j,k}$ are the quantities that are defined at each point of space at each instant, and therefore, the field quantities. They play an important role in Dirac's theory, because they are the ones that present a tensorial character that is analogous to that of the

field quantities in the theories of classical relativistic physics. We shall return to that later on.

As an important example of the application of those general principles, one can consider the case of the z -component of spin. As we have seen, the corresponding operator is $(h/4\pi) i \alpha_1 \alpha_2$. Its proper values – i.e., the results of a possible measurement of S_z – are $\pm h/4\pi$, in accord with the Uhlenbeck-Goudsmit hypothesis.

Upon applying the general principles, one will find, when a choice of the α_i has been made, that the probability of the proper value $+h/4\pi$ for S_z in a state that is represented by a certain wave function ψ will be $\int_D (|\psi_1|^2 + |\psi_3|^2) d\tau$, while the probability of the proper value $-h/4\pi$ will be $\int_D (|\psi_2|^2 + |\psi_4|^2) d\tau$. The sum of those two probabilities is indeed equal to 1, since ψ has been normalized. One can easily recover the value of those probabilities by calculating the mean value of S_z . Indeed, one will have, by definition:

$$\bar{S}_z = \frac{h}{4\pi} \int_D \sum_{k=1}^4 \psi_k^* i \alpha_1 \alpha_2 \psi_k d\tau.$$

Now, with the values that were adopted for the α_i , the rule for matrix multiplication will give:

$$i \alpha_1 \alpha_2 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix},$$

so:

$$\bar{S}_z = \frac{h}{4\pi} \int_D (|\psi_1|^2 - |\psi_2|^2 + |\psi_3|^2 - |\psi_4|^2) d\tau,$$

which agrees with the expressions that were given above for the probabilities of the two possible values $\pm h/4\pi$.

2. THE FIELD QUANTITIES THAT ARE DEFINED BY DIRAC'S THEORY

We have already pointed out the manner by which non-relativistic wave mechanics introduces the field quantities in the form of densities that permit one to reassert the viewpoint of corpuscular physics from that of field physics and which have simple relativistic variances that give a relativistic aspect to all of the formalism. Here however, as in non-relativistic wave mechanics, those field quantities are presented in the form of densities that have an uncertain physical significance from the quantum viewpoint. It is the *spatial* integrals of those quantities (i.e., the mean values or matrix elements) that have a definite physical significance from the quantum viewpoint. However, by contrast, those integrals do not have a relativistic tensorial character (due to just that spatial

integration itself). That curious situation is one aspect of the disagreement that exists between quantum concepts and relativistic concepts.

One can rationally classify the field quantities of Dirac's theory by starting with the following remark: The matrices $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (which play the role of operators that act upon the indices k of the functions ψ_k), when one appends the matrix 1 with four rows and four columns, permit one to define sixteen independent linear, Hermitian matrices by multiplication, which we arrange into the following table:

$$\begin{aligned} \Gamma_0 &= \alpha_4 \\ \Gamma_1 &= \alpha_1 & \Gamma_2 &= \alpha_2 & \Gamma_3 &= \alpha_3 & \Gamma_4 &= 1 \\ \Gamma_{23} &= i\alpha_2 \alpha_3 \alpha_4, & \Gamma_{31} &= i\alpha_3 \alpha_1 \alpha_4, & \Gamma_{12} &= i\alpha_1 \alpha_2 \alpha_4, \\ \Gamma_{14} &= i\alpha_1 \alpha_4, & \Gamma_{24} &= i\alpha_2 \alpha_4, & \Gamma_{34} &= i\alpha_3 \alpha_4, \\ \Gamma_{234} &= i\alpha_2 \alpha_3 \alpha_4 & \Gamma_{314} &= i\alpha_3 \alpha_1 \alpha_4 & \Gamma_{124} &= i\alpha_1 \alpha_2 \alpha_4 & \Gamma_{123} &= i\alpha_1 \alpha_2 \alpha_3 \\ \Gamma_{1234} &= \alpha_1 \alpha_2 \alpha_3 \alpha_4. \end{aligned}$$

In this table, the factors i have been introduced in order to make the matrix products Hermitian, since they would be anti-Hermitian without that factor. We can form sixteen mean-value densities from these sixteen operators that have a simple relativistic variance. Here, we shall specify them upon employing a very customary abbreviation that consists

of writing $\psi^* A \psi$, instead of $\sum_{k=1}^4 \psi_k^* A \psi_k$:

$$(VI.a) \left\{ \begin{array}{llll} & I_1 = \psi^* \alpha_4 \psi & & \\ f_x = -c \psi^* \alpha_1 \psi & f_y = -c \psi^* \alpha_2 \psi & f_z = -c \psi^* \alpha_3 \psi & \rho = \psi^* \psi \\ \mu_x = B \psi^* i \alpha_2 \alpha_3 \alpha_4 \psi & \mu_y = B \psi^* i \alpha_3 \alpha_1 \alpha_4 \psi & \mu_z = B \psi^* i \alpha_1 \alpha_2 \alpha_4 \psi & \\ \pi_x = B \psi^* i \alpha_1 \alpha_4 \psi & \pi_y = B \psi^* i \alpha_2 \alpha_4 \psi & \pi_z = B \psi^* i \alpha_3 \alpha_4 \psi & \\ \sigma_x = \frac{h}{4\pi} \psi^* i \alpha_2 \alpha_3 \psi & \sigma_y = \frac{h}{4\pi} \psi^* i \alpha_3 \alpha_1 \psi & \sigma_z = \frac{h}{4\pi} \psi^* i \alpha_1 \alpha_2 \psi & \sigma_4 = \frac{-h}{4\pi} \psi^* i \alpha_1 \alpha_2 \alpha_3 \psi \\ & I_2 = \psi^* \alpha_1 \alpha_2 \alpha_3 \alpha_4 \psi & & \end{array} \right.$$

The quantity I_1 is invariant under a Lorentz transformation, since that would result from the transformations of ψ_k . Similarly, I_2 is another invariant (or more precisely, a completely-antisymmetric fourth-rank tensor, which amounts to practically the same thing).

The quantities \mathbf{f} and ρ in the second row form the components of space-time quadri-vector. It is the *density-flux* quadri-vector for the probability of presence, which satisfies the continuity equation $\partial \rho / \partial t + \text{div } \mathbf{f} = 0$.

The six quantities μ_x, \dots, π_z form the six distinct components of an antisymmetric second-rank tensor. The vector $\boldsymbol{\mu}$ represents the density of proper magnetic moment for the electron in the state ψ , because upon integrating the components of $\boldsymbol{\mu}$ over that domain, one will get the mean values of the components of the proper magnetic moment

of the electron that were defined above. Similarly, the vector $\boldsymbol{\pi}$ gives the mean density of the proper electric moment, which is the relativistic complement to the mean proper magnetic moment.

The quantities σ_x , σ_y , σ_z , and σ_4 transform like the components of a space-time quadri-vector under a Lorentz transformation (or more precisely, like the components of a completely-antisymmetric third-rank tensor, which amounts to practically the same thing). The components of that quadri-vector are coupled to the invariant I_2 by the relation:

$$\frac{1}{c} \frac{\partial \sigma_4}{\partial t} + \operatorname{div} \boldsymbol{\sigma} = -m_0 c I_2.$$

The quantities σ_x , σ_y , and σ_z are the densities of the mean value for the components of the proper kinetic moment of the electron in the state ψ , and σ_4 is the temporal component that completes the quadri-vector from the relativistic viewpoint. The spin of the electron is then well-defined by a quadri-vector density (and not by an antisymmetric tensor of rank 2), and that is, in fact, what we have said. Later on, we shall see that in the case in which the wave ψ is planar and monochromatic, one can attribute a well-defined velocity \mathbf{v} to the electron, and one will have the relation:

$$\sigma_4 = \frac{1}{c} (\boldsymbol{\sigma} \cdot \mathbf{v})$$

that we encountered before in our relativistic consideration of spin.

Finally, in addition to the sixteen quantities in the preceding table, we can add sixteen more field quantities that define the sixteen components of a second-rank asymmetric tensor that one can attribute to an energy-impulse tensor density for the spinning electron, if one follows the ideas of Costa de Beauregard and van Weysenhoff.

Upon employing the real space-time coordinates, we define those quantities by setting:

$$(VI.b) \quad \left\{ \begin{array}{l} T_{ij} = \frac{hc}{4\pi i} \left[\psi^* \alpha_i \frac{\partial \psi}{\partial x_j} - \frac{\partial \psi^*}{\partial x_j} \alpha_i \psi \right] \quad (i, j = 1, 2, 3), \\ T_{4i} = -\frac{hc}{4\pi i} \left[\psi^* \alpha_i \frac{1}{c} \frac{\partial \psi}{\partial t} - \frac{1}{c} \frac{\partial \psi^*}{\partial t} \alpha_i \psi \right], \\ T_{i4} = -\frac{hc}{4\pi i} \left[\psi^* \frac{\partial \psi}{\partial x_i} - \frac{\partial \psi^*}{\partial x_i} \psi \right], \\ T_{44} = \frac{hc}{4\pi i} \left[\psi^* \frac{1}{c} \frac{\partial \psi}{\partial t} - \frac{1}{c} \frac{\partial \psi^*}{\partial t} \psi \right]. \end{array} \right.$$

That asymmetric tensor answers to the model $\left\| \begin{array}{c|c} \frac{g_i v_k}{c} & c \mathbf{g} \\ \hline \frac{W}{c} \mathbf{v} & W \end{array} \right\|$ that was encountered already

in van Weyssenhoff's theory. T_{44} is the mean value of the energy density. The T_{i4} are the mean values of the components of the impulse density, multiplied by c . The T_{4i} are the mean values of the energy density fluxes along the axes, divided by c . The T_{ij} , for $i, j = 1, 2, 3$ are the mean values of the fluxes of the impulse density along the axes. One will easily confirm those interpretations in the case where ψ is a monochromatic plane wave. The Dirac equations permit one to establish the formula:

$$\sum_{j=1}^4 \frac{\partial T_{ij}}{\partial x_j} = 0 \quad (i = 1, 2, 3, 4).$$

Those four relations express the conservation of the components of the impulse and energy.

One often introduces the symmetric tensor $T'_{ij} = \frac{1}{2}(T_{ij} + T_{ji})$, in place of the asymmetric tensor T_{ij} , and one can verify that it also obeys the conservation relation:

$$\sum_{j=1}^4 \frac{\partial T'_{ij}}{\partial x_j} = \sum_{j=1}^4 \frac{\partial T'_{ji}}{\partial x_j} = 0.$$

Here, as in van Weyssenhoff's theory, the asymmetry of the tensor T_{ij} implies that one does not have $\mathbf{g} \sim W / c^2 \mathbf{v}$, as in the ordinary relativistic mechanics of an electron (without spin). If that were true then the tensor would take the symmetric form:

$$\left\| \begin{array}{c|c} \frac{W}{c^2} v_i v_k & \frac{W}{c} \mathbf{v} \\ \hline \frac{W}{c} \mathbf{v} & W \end{array} \right\|.$$

However, that is not the case in Dirac's theory, except in the case of a monochromatic plane wave, as we shall see much later. One then has $\mathbf{g} \neq \mathbf{v}$.

3. MONOCHROMATIC PLANE WAVES IN DIRAC'S THEORY

From the fundamental ideas of wave mechanics, a uniform, rectilinear motion of an electron with energy W and quantity of motion \mathbf{p} must correspond to a monochromatic plane wave, which will have the components:

$$\psi_k = a_k e^{\frac{2\pi i}{h}(Wt - \mathbf{p} \cdot \mathbf{r})} \quad (k = 1, 2, 3, 4)$$

in Dirac's theory. Upon substituting that form in the four Dirac equations, one will get four linear, homogeneous, algebraic relations between the four amplitudes a_k . In order for those equations to admit solutions for the a_k that are not all zero, their determinant must be zero. If one performs the calculation then one will find the condition:

$$\frac{W^2}{c^2} = p_x^2 + p_y^2 + p_z^2 + m_0^2 c^2;$$

i.e., the relation that couples the energy and quantity of motion of a free corpuscle of proper mass m_0 in relativistic theory. If that condition is satisfied then not only will the determinant of the equations in a_k be zero, but also the first-order minors, in such a way that two of the a_k will be arbitrary, while the other two will be expressed in terms of them. If we set $a_3 = c_1$ and $a_4 = c_2$ arbitrarily then we will get:

$$a_1 = -\frac{p_z c_1 + (p_x + ip_y) c_2}{W/c + m_0 c}, \quad a_2 = -\frac{(p_x - ip_y) c_1 - p_z c_2}{W/c + m_0 c},$$

$$a_3 = c_1, \quad a_4 = c_2.$$

In order to interpret that result, suppose that the z -axis is taken to be in the direction of motion (viz., the direction of the vector \mathbf{p}). That is no loss of generality, since one knows, in principle, how to transform the ψ_k for a rotation of the spatial axes. One will then have $p_x = p_y = 0$, and:

$$\psi_1 = -\frac{P_z}{\Delta} c_1 P, \quad \psi_2 = \frac{P_z}{\Delta} c_2 P, \quad \psi_3 = c_1 P, \quad \psi_4 = c_2 P,$$

with:

$$\Delta = \frac{W}{c} + m_0 c, \quad P = e^{\frac{2\pi i}{h}(Wt - p_z z)}.$$

One can state that result in the following fashion: Any monochromatic plane wave that corresponds to a uniform, rectilinear motion (with positive energy) of a Dirac corpuscle can be considered to be the superposition of two waves $\psi^{(g)}$ and $\psi^{(d)}$ according to the formula ([†]):

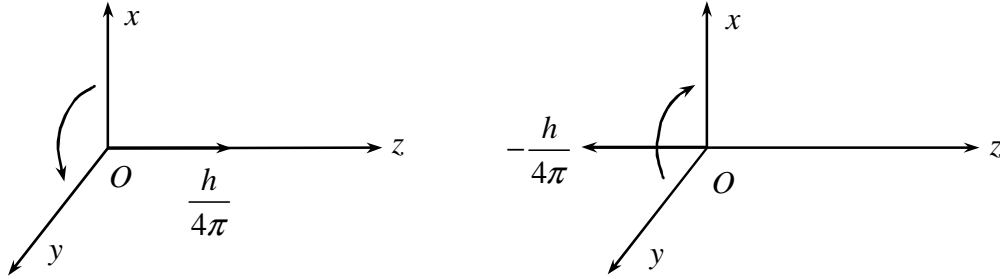
$$\psi = c_1 \psi^{(g)} + c_2 \psi^{(d)},$$

where $\psi^{(g)}$ and $\psi^{(d)}$ have the following components:

$$\psi_1^{(g)} = -\frac{P_z}{\Delta} P, \quad \psi_3^{(g)} = P, \quad \psi_2^{(g)} = \psi_4^{(g)} = 0; \quad \psi_1^{(d)} = \psi_2^{(d)} = 0, \quad \psi_3^{(d)} = \frac{P_z}{\Delta} P, \quad \psi_4^{(d)} = P.$$

([†]) Translator: The choice of notation comes from the French $g = gauche$ = left and $d = droit$ = right, as will become clear below.

If one recalls that the probability of the two possible values $\pm h / 4\pi$ of the z -component of spin are proportional to $|\psi_1|^2 + |\psi_3|^2$ and $|\psi_2|^2 + |\psi_4|^2$, resp., then one will see that $\psi^{(g)}$ will correspond to spin $+ h / 4\pi$, and $\psi^{(d)}$, to spin $- h / 4\pi$. Upon associating a vector that is carried by Oz with a rotation in the plane xOy by the usual rule, one will make $\psi^{(g)}$ correspond to a “levogyrous” or *left-circular* rotation and $\psi^{(d)}$ will correspond to a “dextrogyrous” or *right-circular* rotation, according to the schema below:



Any monochromatic plane wave (with positive energy) will then be the sum of two monochromatic plane waves that each correspond to a proper value of the spin component in the direction of propagation, and the proportion between the two waves and their phase difference in that superposition will be given by the values of the two complex components C_1 and C_2 .

One sees in the preceding formulas that ψ_1 and ψ_2 will become negligible in comparison to ψ_3 and ψ_4 when $p_z c$ is small in comparison to $m_0 c^2$; i.e., in the approximation of Newtonian mechanics ($v \ll c$). In the opposite limiting case of a velocity that is close to c , $m_0 c^2$ will be negligible in comparison to W , and $|\mathbf{p}|$ will be roughly equal to W / c ; one will then have:

$$-\psi_1^{(g)} = \psi_3^{(g)} = P, \quad \psi_2^{(g)} = \psi_4^{(g)} = 0; \quad \psi_2^{(d)} = \psi_4^{(d)} = P, \quad \psi_1^{(d)} = \psi_3^{(d)} = 0.$$

It is very interesting to calculate the sixteen mean densities that are defined by the Table (VI.a) for the case of a monochromatic plane wave. We first remark that the constants C_1 and C_2 are coupled by the normalization relation for the wave ψ . In order to avoid the introduction of formulas that are somewhat complicated and relate to continuous spectra, we shall employ the often-used trick of supposing that the corpuscle is enclosed in a box of volume \mathcal{V} whose dimensions are very large with respect to the wave length. We can then write the normalization condition as:

$$(VI.c) \quad \int_D \sum_{k=1}^4 \psi_k^* \psi_k d\tau = (|C_1|^2 + |C_2|^2) \left[1 + \frac{p^2}{\Delta^2} \right] \mathcal{V} = 1,$$

which can also be written:

$$(|C_1|^2 + |C_2|^2) \frac{2W/c}{\Delta} \mathcal{V} = 1.$$

Now calculate the sixteen mean densities in the Table (VI.a), while taking the preceding relations into account. One first finds that:

$$I_1 = |\psi_3|^2 + |\psi_4|^2 - |\psi_1|^2 - |\psi_2|^2 = \frac{m_0 c^2}{W} \cdot \frac{1}{\mathcal{V}}.$$

If we introduce the velocity βc of the corresponding rectilinear motion into that expression such that $W = \frac{m_0 c^2}{\sqrt{1-\beta^2}}$ then we will have:

$$I_1 = \sqrt{1-\beta^2} \frac{1}{\mathcal{V}}, \quad \int_D I_1 d\tau = \sqrt{1-\beta^2},$$

and we will see that the invariant I_1 is closely linked with the Lorentz contraction. The contraction factor $\frac{m_0 c^2}{W}$ plays an important role here.

$$I_2 = i [\psi_1^* \psi_3 + \psi_2^* \psi_4 - \psi_3^* \psi_1 - \psi_4^* \psi_2] = 0.$$

That invariant is then zero for a monochromatic plane wave. An analogous calculation will provide the components of the “flux-density” quadri-vector:

$$f_x = f_y = 0, \quad f_z = (|C_1|^2 + |C_2|^2) \frac{2pc}{\Delta}, \quad \rho = (|C_1|^2 + |C_2|^2) \frac{2W/c}{\Delta}.$$

Upon comparing the expressions for f_z and ρ and recalling the fact that $\rho = Wv / c^2$, one will have:

$$f_x = f_y = 0, \quad f_z = \rho \frac{\rho c^2}{W} = \rho v.$$

There is then a probability flux in the direction of propagation of the wave ψ with a velocity of v ; that is what one should expect.

We now go on to the calculation of the “spin density” quadri-vector:

$$\begin{aligned} \sigma_x &= \frac{h}{4\pi} (c_1^* c_2 + c_2^* c_1) \frac{2m_0 c}{\Delta}, & \sigma_y &= \frac{h}{4\pi} i (c_1^* c_2 - c_2^* c_1) \frac{2m_0 c}{\Delta}, \\ \sigma_z &= \frac{h}{4\pi} (|c_1|^2 - |c_2|^2) \frac{2W/c}{\Delta}, & \sigma_4 &= \frac{h}{4\pi} (|c_1|^2 - |c_2|^2) \frac{2p}{\Delta}. \end{aligned}$$

The normalization relation for ψ permits one to write:

$$\sigma_z = \frac{1}{V} \cdot \frac{h}{4\pi} \cdot \frac{|C_1|^2 - |C_2|^2}{|C_1|^2 + |C_2|^2}, \quad S_z = \int_D \sigma_z d\tau = \frac{h}{4\pi} \cdot \frac{|C_1|^2 - |C_2|^2}{|C_1|^2 + |C_2|^2}.$$

Upon comparing the expressions for σ_z and σ_4 , one will infer that:

$$\sigma_4 = \sigma_z \frac{pc}{W} = \frac{1}{c} (\boldsymbol{\sigma} \cdot \mathbf{v}),$$

which is the relation that previously predicted.

Finally, we give the components of the mean densities of the proper electric and magnetic moments for the monochromatic plane wave:

$$(VI.d) \quad \begin{cases} \mu_x = -B(C_1^* C_2 + C_2^* C_1) \frac{2W/c}{\Delta}, & \mu_y = Bi(C_1^* C_2 - C_2^* C_1) \frac{2W/c}{\Delta}, & \mu_z = B(|C_2|^2 - |C_1|^2) \frac{2m_0 c}{\Delta}, \\ \pi_x = Bi(C_2^* C_1 - C_1^* C_2) \frac{2p}{\Delta}, & \pi_y = B(C_2^* C_1 + C_1^* C_2) \frac{2p}{\Delta}, & \pi_z = 0. \end{cases}$$

One easily verifies from these formulas that:

$$\boldsymbol{\pi} = [\boldsymbol{\mu} \times \frac{1}{c} \mathbf{v}], \quad (\boldsymbol{\pi} \cdot \mathbf{v}) = 0,$$

which are formulas that were likewise predicted before. We see that the vector $\boldsymbol{\pi}$ is always transversal with respect to the direction of propagation, while $\boldsymbol{\mu}$ can make an arbitrary angle with it, but $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ will always be mutually perpendicular.

When the velocity v tends to c , the vector $\boldsymbol{\sigma}$ will lie along the direction of propagation. On the contrary, $\boldsymbol{\mu}$ will tend to something that is perpendicular to that direction, so in the limit, the vectors $\boldsymbol{\mu}$ and $\boldsymbol{\pi}$ will both be in the wave plane, normal to each other, and equal in magnitude, like the electromagnetic field of a light wave. Hence, the vectors $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, whose directions coincide in the proper system of the particle, will form a right angle for an observer that sees the particle pass with a velocity that is very close to c .

We just studied monochromatic plane waves with positive energy W . In Dirac's theory, there also exist plane waves with negative energies that play a very special role in the theory. We shall make a special study of them later on

4. THE DENSITY-CURRENT QUADRI-VECTOR AND ITS DECOMPOSITION

We defined the quantities ρ and \mathbf{f} for the electron above. Upon multiplying them by the electric charge of the electron $-e$, one will get a quadri-vector whose temporal component gives the mean density $-e |\psi|^2 = \delta$ of the electricity for the electron in the state ψ , and whose spatial components:

$$j_i = -e f_i = e c \sum_{k=1}^4 \psi_k^* \alpha_i \psi_k \quad (i = 1, 2, 3)$$

will give the components of the corresponding current density.

We shall first show how one can decompose the current \mathbf{j} into two parts, and the result that is obtained, which is called the *Gordon decomposition*, will show us that the current \mathbf{j} will not reduce to a convection current that is due to the displacement of the charged corpuscle, as a result of the existence of the proper magnetic moment and the proper electric moment that is implied by the form itself of the Dirac equations.

In order to make that decomposition, we start from the Dirac equations in the absence of a field:

$$\frac{1}{c} \frac{\partial \psi_k}{\partial t} = \frac{\partial}{\partial x} \alpha_1 \psi_k + \frac{\partial}{\partial y} \alpha_2 \psi_k + \frac{\partial}{\partial z} \alpha_3 \psi_k + \kappa m_0 c \alpha_4 \psi_k,$$

and from the Hermitian character of the α_i , the conjugate equations will permit one to write:

$$\frac{1}{c} \frac{\partial \psi_k^*}{\partial t} = \frac{\partial}{\partial x} \psi_k' \alpha_1 + \frac{\partial}{\partial y} \psi_k' \alpha_2 + \frac{\partial}{\partial z} \psi_k' \alpha_3 + \kappa m_0 c \psi_k^* \alpha_4,$$

with the convention that:

$$\psi_k^* \alpha_i = \sum_{l=1}^4 \psi_l^* (\alpha_i)_{lk}.$$

Multiply the first equation on the left by $\psi_k^* \alpha_1 \alpha_4$ and the second one on the right by $\alpha_1 \alpha_4 \psi_k$, and then add them and sum over k . One will easily get, in symbolic notation:

$$\psi^* \alpha_1 \psi = \frac{h}{4\pi i m_0 c} \left[\frac{1}{c} \frac{\partial}{\partial t} (\psi^* \alpha_1 \alpha_4 \psi) + \psi^* \alpha_4 \frac{\partial \psi}{\partial x} - \frac{\partial \psi^*}{\partial x} \alpha_4 \psi + \frac{\partial}{\partial y} (\psi^* \alpha_1 \alpha_2 \alpha_3 \psi) - \frac{\partial}{\partial z} (\psi^* \alpha_3 \alpha_1 \alpha_4 \psi) \right],$$

so upon multiplying this by $e c$ and taking into account the definitions that were assumed, one will easily deduce that:

$$j_x = \frac{\partial \pi_x}{\partial t} + c (\text{rot } \boldsymbol{\mu})_x + \frac{eh}{4\pi m_0 i} \left(\psi^* \alpha_4 \frac{\partial \psi}{\partial x} - \frac{\partial \psi^*}{\partial x} \alpha_4 \psi \right).$$

One similarly proves the two relations that are deduced by circular permutation of x, y, z , and one will conclude that:

$$\mathbf{j} = \mathbf{j}_1 + \mathbf{j}_2,$$

with:

$$\mathbf{j}_1 = \frac{eh}{4\pi m_0 i} (\psi^* \alpha_4 \text{grad } \psi - \text{grad } \psi^* \alpha_4 \psi),$$

$$\mathbf{j}_2 = \frac{\partial \rho}{\partial t} + c \text{rot } \boldsymbol{\mu}.$$

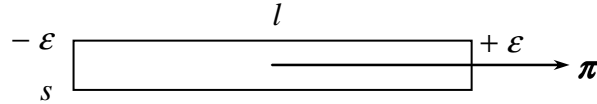
The term \mathbf{j}_1 is easy to interpret when one considers the case in which the wave ψ is monochromatic and planar; i.e., when one attributes a uniform, rectilinear motion to the electron with a velocity of $\mathbf{v} = (c^2 / W) \mathbf{p}$. One will then have:

$$\mathbf{j}_1 = \frac{eh}{4\pi m_0 i} 2 \left(-\frac{2\pi i}{h} \mathbf{p} \right) \psi^* \alpha_4 \psi = -\rho e \frac{\mathbf{p}}{m_0} \frac{m_0 c^2}{W} = -\rho e \frac{\mathbf{p} c^2}{W} = (-\rho e) \mathbf{v}.$$

\mathbf{j}_1 is then simply the mean density of the convection current that is due to the *collective motion* of the electric charge.

The interpretation of the term \mathbf{j}_2 is more interesting, but it requires some preliminary remarks.

First consider a polarized dielectric medium, and in that medium, consider a cylindrical element of length l and section s whose axis coincides with the direction of the polarization vector.



One of the bases carries the charge $+\varepsilon$, and the other one, the charge $-\varepsilon$. The electric moment is εl , which corresponds to the density $\pi = \frac{\varepsilon l}{ls} = \varepsilon / s$. If the first charge becomes $\varepsilon + d\varepsilon$ and the second one, $-(\varepsilon + d\varepsilon)$, then the density will be increased by $d\varepsilon / s$, and if that increase takes place during a time interval dt then one can consider it to be due to the transport of the charge $d\varepsilon$ from the second base to the first one, so it will be coupled with a current that is equal to $d\varepsilon / dt$; i.e., a current density of $\frac{1}{s} \frac{d\varepsilon}{dt}$. Finally, one sees that the increase of π is equivalent to the existence of a current density that is equal to $d\pi / dt$.

On the other hand, consider a magnet with a magnetic moment of \mathfrak{M} . As one knows, the magnetic field that is created by that magnet is derived from the vector potential $-(\mathfrak{M} \cdot \text{grad } 1/r)$, but it is also derived from the vector potential $\mathbf{A} = -(\mathfrak{M} \times \text{grad } 1/r)$, as one easily sees by proving that:

$$-\text{grad} \left[\mathfrak{M} \cdot \text{grad} \frac{1}{r} \right] = \text{rot} \left[\mathfrak{M} \times \text{grad} \frac{1}{r} \right].$$

The field that is created outside by a distribution of magnetism of magnetization $\boldsymbol{\mu}$ that occupies a volume of \mathcal{V} is derived from the vector potential:

$$\mathbf{A} = - \int_{\nu} \left[\boldsymbol{\mu} \times \text{grad} \frac{1}{r} \right] d\tau = \int_{\nu} \frac{\text{rot } \boldsymbol{\mu}}{r} d\tau,$$

in which the latter expression is provided by an integration by parts. Upon comparing it with the classical formula for a retarded potential:

$$\mathbf{A} = \frac{1}{c} \int \frac{\mathbf{j}}{r} d\tau,$$

one will then see that the magnetization $\boldsymbol{\mu}$ is equivalent to a current density that equals $c \text{ rot } \boldsymbol{\mu}$.

In summary, we have recovered the classical result:

If a region in space is filled with a distribution of electric moments that is defined at each point by a certain polarization $\boldsymbol{\pi}$ and a distribution of magnetic moments that is defined at each point by a certain magnetization $\boldsymbol{\mu}$ then that region of space will be filled with microscopic currents whose density is given by the vector $\frac{\partial \boldsymbol{\pi}}{\partial t} + c \text{ rot } \boldsymbol{\mu}$.

If we then return to the Gordon decomposition then we will see that the vector \mathbf{j} has its origin in the existence of proper dielectric and magnetic moments for the electron.

The region that is occupied by the electron wave ψ can be compared (without taking the comparison too literally) to the polarized and magnetized medium of classical theory. The electricity will possess a global convective motion in it that is characterized by the velocity $\mathbf{v} = (c^2 / M) \mathbf{p}$ in the case where the collective motion is uniform and rectilinear. It is superimposed with a “fine-grained” motion of the electricity that corresponds to the polarization and magnetization.

Upon regarding things roughly, one will see that one must distinguish between the global “fine-grained” velocity \mathbf{u} whose components corresponds to the operators $-c\alpha_1$, $-c\alpha_2$, $-c\alpha_3$, and the collective translational velocity \mathbf{v} whose components corresponds to

the operators $-\frac{h}{2\pi i m_0} \alpha_4 \frac{\partial}{\partial x}$, $-\frac{h}{2\pi i m_0} \alpha_4 \frac{\partial}{\partial y}$, $-\frac{h}{2\pi i m_0} \alpha_4 \frac{\partial}{\partial z}$. The former serves to

define the total “fine-grained” current \mathbf{j} , while the latter defines the convection current \mathbf{j}_1 . The difference $\mathbf{j} - \mathbf{j}_1 = \mathbf{j}_2$ represents the current that is due to the effects of polarization and magnetization. One will find a detailed study of currents in a polarized medium in the book of R. Becker on the electron ⁽¹⁾. The analogy can be extended, but one must not forget that here we have only a mean representation of the currents, polarizations, and magnetizations that are attached to a single particle.

The distinction between \mathbf{u} and \mathbf{v} also makes it comprehensible that the energy-impulse tensor T_{ij} is not symmetric. In reality, it has the matrix:

⁽¹⁾ *Théorie des Electrons*, Alcan, Paris, 1936; pp. 125, *et seq.*

$$\left\| \begin{array}{c|c} g^j u^i & c \mathbf{g} \\ \hline \frac{W}{c} \mathbf{u} & W \end{array} \right\|$$

(in which \mathbf{u} is the total fine-grained velocity that we just defined), and not:

$$\left\| \begin{array}{c|c} g^j v^i & c \mathbf{g} \\ \hline \frac{W}{c} \mathbf{v} & W \end{array} \right\|,$$

and since it is \mathbf{v} , not \mathbf{u} , that is proportional to the quantity of motion, one will see why T_{ij} is not symmetric. The form itself of formulas (IV.b) for the T_{ij} indeed shows that it is composed of densities of quantities of motion and *components of \mathbf{u}* .

We have obtained the formula:

$$\mathbf{j} = \frac{eh}{4\pi i m_0} (\psi^* \alpha_4 \text{grad } \psi - \text{grad } \psi^* \alpha_4 \psi) + \frac{\partial \boldsymbol{\pi}}{\partial t} + c \text{rot } \boldsymbol{\mu}.$$

With the aid of the relativistic notations, it can be written:

$$j_i = \frac{eh}{4\pi i m_0} \left(\psi^* \alpha_4 \frac{1}{c} \frac{\partial \psi}{\partial t} - \frac{1}{c} \frac{\partial \psi^*}{\partial t} \alpha_4 \psi \right) + c \sum_j \frac{\partial m^{ij}}{\partial x_j} \quad (i = 1, 2, 3),$$

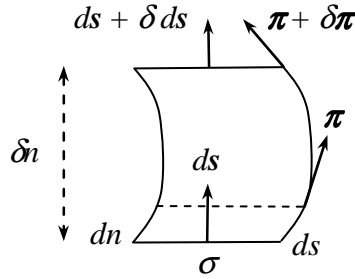
in which m^{ij} is the antisymmetric tensor $(\boldsymbol{\pi}, \boldsymbol{\mu})$. Since the fourth component of the current density is $j_4 = \delta c$, upon applying the preceding formula to the case of $i = 4$, one will likewise find that:

$$\delta = \delta_1 + \delta_2 = \frac{-eh}{4\pi i m_0} \left(\psi^* \alpha_4 \frac{1}{c} \frac{\partial \psi}{\partial t} - \frac{1}{c} \frac{\partial \psi^*}{\partial t} \alpha_4 \psi \right) - \text{div } \boldsymbol{\pi}.$$

The first term corresponds to the density that would exist if the spin did not exist. The second one – namely, $\text{div } \boldsymbol{\pi}$ – corresponds to the existence of spin, because one knows that in an electrically-polarized medium there exists a microscopic density that is equal to $-\text{div } \boldsymbol{\pi}$. We recover the same picture for the fourth component of the current density that we had for the first three [viz., consideration of a microscopic charge

$$\frac{\partial}{\partial t} (-\text{div } \boldsymbol{\pi}) + \text{div} \left(\frac{\partial \boldsymbol{\pi}}{\partial t} + \text{rot } \boldsymbol{\mu} \right) = 0].$$

In order to show the meaning of $-\text{div } \boldsymbol{\pi}$, one can consider an element of a force tube that corresponds to the polarization vector in the polarized medium. On the lower section of that element, the density, transversal section, and polarization will have the values σ , ds , and $\boldsymbol{\pi}$, resp. On the upper section, those values will be increased by $\delta\sigma$, δds , and $\delta\boldsymbol{\pi}$, resp.



One has:

$$-\varepsilon = \sigma ds, \quad \pi = \frac{\varepsilon dn}{dn ds} = \frac{\varepsilon}{ds} = -\sigma.$$

The flux-divergence theorem gives:

$$\iint \pi_n ds = (\boldsymbol{\pi} + \delta \boldsymbol{\pi})(ds + \delta ds) - \boldsymbol{\pi} \delta s \approx \delta(\boldsymbol{\pi} \cdot ds) = \delta(\boldsymbol{\sigma} \cdot ds) = \iiint \operatorname{div} \boldsymbol{\pi} dv;$$

hence:

$$\operatorname{div} \boldsymbol{\pi} = -\frac{\delta(\boldsymbol{\sigma} ds)}{\delta n ds}.$$

The charge on the volume element dv is then $\int_0^{\delta n} \frac{d(\boldsymbol{\sigma} ds)}{dn} dn = \delta(\boldsymbol{\sigma} ds)$, and the charge

density in the element is $\frac{1}{\delta n ds} d(\boldsymbol{\sigma} ds) = -\operatorname{div} \boldsymbol{\pi}$. Q. E. D.

CHAPTER VII

NEGATIVE-ENERGY SOLUTIONS IN DIRAC'S THEORY

1. THE NEGATIVE-ENERGY PLANE WAVE

We just studied solutions of the Dirac equation that represent monochromatic plane waves of positive energy W , and we found that in order to have a solution of the form:

$$\psi_k = a_k e^{\frac{2\pi i}{h}(Wt - \mathbf{p}\cdot\mathbf{r})},$$

we must have the following relation between W and \mathbf{p} :

$$(VII.a) \quad \frac{W^2}{c^2} = p^2 + m_0^2 c^2 = p_x^2 + p_y^2 + p_z^2 + m_0^2 c^2.$$

Upon setting:

$$W = +c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2},$$

we found that:

$$a_1 = -\frac{p_z A + (p_x + ip_y)B}{W/c + m_0 c}, \quad a_2 = -\frac{p_z B - (p_x - ip_y)A}{W/c + m_0 c}, \quad a_3 = A, \quad a_4 = B,$$

in which A and B are arbitrary constants.

We can also satisfy the condition (VII.a) by setting:

$$W = -c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}.$$

We will then get the following solution:

$$a_1 = C, \quad a_2 = D, \quad a_3 = \frac{p_z C + (p_x + ip_y)D}{m_0 c - W/c}, \quad a_4 = \frac{(p_x - ip_y)C - p_z D}{m_0 c - W/c},$$

in which C and D are arbitrary constants.

We shall now slightly modify the notations that are currently employed. For given values of p_x, p_y, p_z , we let W denote the positive quantity $\sqrt{m_0^2 c^2 + p^2}$, and in order to account for the second solution, we say that we have to consider both the wave of energy W and the wave of energy $-W$. One must then change W into $-W$ in the latter formulas.

Briefly, for given values of p_x, p_y, p_z , with:

$$W = +c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2},$$

we will have to consider the monochromatic plane wave:

$$\psi_k = a_k e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)},$$

with:

$$(VII.b) \quad a_1 = -\frac{p_z A + (p_x + ip_y)B}{W/c + m_0 c}, \quad a_2 = \frac{p_z B - (p_x - ip_y)A}{W/c + m_0 c}, \quad a_3 = A, \quad a_4 = B,$$

and the monochromatic plane wave with negative energy – W that is defined by:

$$\psi_k = b_k e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)},$$

with:

$$(VII.c) \quad b_1 = C, \quad b_2 = D, \quad b_3 = \frac{p_z C + (p_x + ip_y)D}{W/c + m_0 c}, \quad b_4 = \frac{p_z D - (p_x - ip_y)C}{W/c + m_0 c}.$$

Let us compare those two waves.

For the positive-energy wave, we already know that the components ψ_3 and ψ_4 , which correspond in some way to the mass m_0 , outweigh the waves ψ_1 and ψ_2 , which correspond to the mass $-m_0$. The waves ψ_1 and ψ_2 will be zero for zero velocity, and will become important only for velocities that are close to c .

Consider the invariant:

$$I_1 = \psi^* \alpha_4 \psi = |\psi_1|^2 + |\psi_2|^2 - |\psi_3|^2 - |\psi_4|^2 = \frac{\sqrt{1-\beta^2}}{\nu}.$$

It is always positive and tends to zero when ν tends to c .

The conclusions are the opposite ones for the negative-energy wave. The waves ψ_1 and ψ_2 will predominate, and if the electron is at rest (i.e., if $p_x = p_y = p_z = 0$) then one will have:

$$\psi_1 = C e^{\frac{-2\pi i}{h} m_0 c^2 t}, \quad \psi_2 = D e^{\frac{-2\pi i}{h} m_0 c^2 t}, \quad \psi_3 = \psi_4 = 0.$$

The waves ψ_3 and ψ_4 will take on more importance for increasing velocities, but it is only for the limiting case of $\nu = c$ that one will have $|\psi_1|^2 + |\psi_2|^2 = |\psi_3|^2 + |\psi_4|^2$, in such a way that one will then have $I_1 \leq 0$, and equality will refer to the limiting case $\nu = c$. It will be the waves of negative proper mass that dominate. One can associate proper mass to the operator $m_0 \alpha_4$, so $I_1 \cdot m_0$ will be the mean-value density. The mean value that corresponds to it is $m_0 \sqrt{1-\beta^2}$ for positive-energy waves and $-m_0 \sqrt{1-\beta^2}$ for negative-energy waves.

The existence of negative-energy states in Dirac's theory constitutes a strange situation. The properties of an electron in such a state will be extraordinary. When placed in an electric field \mathbf{h} , it will accelerate in the opposite sense to the force $-e \mathbf{h}$. One will then increase its velocity while decreasing its energy. Its velocity will be in the

opposite sense to its quantity of motion ⁽¹⁾, etc. One never observes corpuscles that have those paradoxical properties. There is then a problem with Dirac's theory.

One can believe that this difficulty exists already in Einsteinian relativistic dynamics, because the relation:

$$\frac{W^2}{c^2} = m_0 c^2 + p_x^2 + p_y^2 + p_z^2$$

is valid in that theory, and upon taking the square root, one will find that $W = \pm c\sqrt{m_0^2 c^2 + p^2}$ with the double sign \pm . However, it is easy to discard that complication in the present case.

Indeed, the possible values for W are found in the disjoint intervals $(-\infty, -m_0 c^2)$ and $(m_0 c^2, +\infty)$, while the interval $(-m_0 c^2, m_0 c^2)$ does not correspond to any possible value for W . Now, in the old dynamics (relativistic or not), variations take place in a continuous fashion. Hence, if all of the electrons have a positive energy that is found between $m_0 c^2$ and $+\infty$ to begin with then the same thing will always be true after that. No negative value of W that is found between $-\infty, -m_0 c^2$ will appear, since the interval $(-m_0 c^2, m_0 c^2)$ cannot be crossed by a continuous variation of W . The objection that relates to negative energies is then found to be irrelevant in Einsteinian dynamics.

The same thing is not true in quantum mechanics, because, in principle, it admits the existence of sharp transitions between states whose energy will differ by a finite quantity, which will prevent one from discarding *a priori* the passage from a domain of positive energies to one of negative energies. What is more, one can imagine some simple examples in which the transitions of that type are found to be realized.

In an article in 1929 [Zeit. Phys. **53** (1929), pp. 157], O. Klein was the first to point out an example of a transition, which, without being a passage from a state of positive energy to a state of negative energy, strictly speaking, is nonetheless equivalent to one. He considered a planar surface S that separated a region (I) in which the potentials are zero from a region (II) in which a constant, negative scalar potential prevails, in such a way that an electron would possess the potential energy $U = -eV > 0$ in region (II). A Dirac electric wave from region (I) falls normally upon the separating surface; that wave is assumed to be monochromatic, planar, and have positive energy W . Everything comes down to calculating the waves that are reflected and transmitted by the separating surface.

$$\begin{array}{c} U_1 = 0 \quad \Big| \quad U_2 = -eV > 0 \\ \text{(I)} \quad \Big| \quad \text{(II)} \\ S \end{array}$$

One shows that in order to perform the calculation, one must express the continuity of each of the four ψ_k upon crossing the surface S ; one must say that:

$$\psi_k \text{ incident} + \psi_k \text{ reflected} = \psi_k \text{ transmitted.}$$

⁽¹⁾ By virtue of the relation $\bar{v} = \partial W / \partial p$, which expresses the idea that the velocity of the corpuscle = group velocity.

Naturally, the reflected and transmitted waves will correspond to the same energy W as the incident wave: The phenomenon is conservative.

However, Klein showed the following result: For:

$$0 < U < W - m_0 c^2,$$

there will be both reflection and transmission, and the transmitted wave, like the reflected wave, will have the normal character of a positive-energy wave. For $W - m_0 c^2 < U < W + m_0 c^2$, there will be total reflection, with a vanishing wave in the second medium. Finally, for $U > W + m_0 c^2$, one will once more find a transmitted wave that crosses S , but that wave will be a sort of negative-energy wave, which is what constitutes the “Klein paradox.” Obviously, its total energy W is positive, but that is what one calls the “energy of a non-potential nature”; i.e., $W - U$ is negative and less than $-m_0 c^2$ in region (II), whereas it must always be greater than $m_0 c^2$ in Einsteinian dynamics. The transmitted wave in the medium (II), in which the scalar potential V rules, will then be analogous to a negative-energy wave in the absence of a potential, and will possess the same paradoxical properties. One must say that there is a certain probability that an incident electron will penetrate into region (II) and pass into that strange state. That is the *Klein paradox*.

The case that Klein envisioned is very schematic. Other authors, notably Gerard Petiau in France, have advanced the study of that type of problem. The essential result seems to be the following one: Whenever the potential energy of the electron submits to a variation that is equal to at least $m_0 c^2$ over a distance of order $h / m_0 c$, there will be the possibility of passing to states of negative energy. Here, one will see that role that is played by the “Compton wave length” $h / m_0 c = 2.3 \times 10^{-10}$ cm, which we will encounter again later on.

2. INCOMPLETE CHARACTER OF THE SYSTEM OF POSITIVE-ENERGY WAVES

Let us return to certain peculiarities of non-relativistic wave mechanics. In that theory, the equation of propagation $\frac{h}{2\pi i} \frac{\partial \psi}{\partial t} = H\psi$ has order one with respect to time.

The solution will then be determined completely when one gives its initial form $\psi(x, y, z, 0)$. Consider the case in which a field is absent and the equation of propagation takes the simple form:

$$\Delta \psi = \frac{4\pi i m}{h} \frac{\partial \psi}{\partial t}.$$

It has the monochromatic plane wave for a solution:

$$\psi(x, y, z, t) = a e^{\frac{2\pi i}{h}(Et - p_x x - p_y y - p_z z)},$$

with:

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \quad (\text{no sign ambiguity}).$$

Suppose that we give the initial form of the wave function $\psi(x, y, z, 0) = F(x, y, z)$ and consider the development of $F(x, y, z)$ in a Fourier integral:

$$F(x, y, z) = \int \int \int_{-\infty}^{+\infty} g(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dp_x dp_y dp_z.$$

The coefficients g can be calculated by starting with a given F from the formula:

$$g(p_x, p_y, p_z) = \frac{1}{h^2} \int \int \int_{-\infty}^{+\infty} F(x, y, z) e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dx dy dz.$$

I then say that the function:

$$\psi(x, y, z, t) = \int \int \int_{-\infty}^{+\infty} g(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(Et - p_x x + p_y y + p_z z)} dp_x dp_y dp_z$$

is the solution that admits F for its initial form. Indeed:

1. $\psi(x, y, z, t)$ satisfies the equation of propagation, since it is a sum of monochromatic plane solutions of that linear equation.

2. It obviously reduces to $F(x, y, z)$ at the instant $t = 0$.

We then conclude that in non-relativistic wave mechanics, the system of monochromatic plane waves constitutes a “complete” system in the sense that any solution to the wave equation can be represented as a superposition of monochromatic plane waves.

We now pass on to Dirac’s theory and seek to transpose the preceding argument into it. There, we will be dealing with four simultaneous first-order equations in x, y, z, t in the four ψ_k . The four wave functions will then be determined entirely when one is given the initial form $\psi(x, y, z, 0)$. As always, take the case of a zero field, and once more demand to know whether it is possible to represent no particular solution by a superposition of monochromatic planes. A solution is obviously defined by being given the four:

$$\psi_k(x, y, z, 0) = F_k(x, y, z),$$

which we assume to be developed into Fourier integrals of the form:

$$F_k(x, y, z) = \int \int \int_{-\infty}^{+\infty} g_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dp_x dp_y dp_z,$$

in which the g_k are given by the inversion formulas:

$$g_k(p_x, p_y, p_z) = \frac{1}{h^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_k(x, y, z) e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dx dy dz.$$

We seek to represent the solution that corresponds to a given initial ψ_k by a superposition of monochromatic plane waves that contain only positive-energy waves. In order to do that, set:

$$\psi_k(x, y, z, t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(Wt - p_x x + p_y y + p_z z)} dp_x dp_y dp_z,$$

with $W = +c\sqrt{m_0^2 c^2 + p^2}$. We will indeed have a solution to the wave equation with that, but in order for it to have the initial form $F(x, y, z)$, one must have:

$$a_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z) \quad (k = 1, 2, 3, 4),$$

in which the g_k are known. However, we know that of the four a_k , only two of them are arbitrary, and that will show us that we cannot satisfy the preceding conditions, in general. The positive-energy monochromatic plane waves do not form a complete system in the present case then.

Now consider the negative-energy monochromatic plane waves, along with the positive-energy monochromatic plane waves. We will then obtain a complete system. Indeed, if we set:

$$\begin{aligned} \psi_k(x, y, z, t) = & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(Wt - \mathbf{p} \cdot \mathbf{r})} dp_x dp_y dp_z \\ & + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} b_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(-Wt - \mathbf{p} \cdot \mathbf{r})} dp_x dp_y dp_z, \end{aligned}$$

with $W = +c\sqrt{m_0^2 c^2 + p^2}$, as always, then we will have a solution of the equations of propagation, since it is linear, and the $\psi_k(x, y, z, 0)$ will coincide with the $F_k(x, y, z)$ if one has:

$$a_k(p_x, p_y, p_z) + b_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z) \quad (k = 1, 2, 3, 4).$$

This will always be compatible with the difference of the conditions obtained above, because four of the eight a_k and b_k are arbitrary.

The preceding conditions are written explicitly as:

$$\begin{aligned} -\frac{p_z A + (p_x + ip_y)B}{W/c + m_0 c} + C = g_1(p_x, p_y, p_z), \quad \frac{p_z B - (p_x - ip_y)A}{W/c + m_0 c} + D = g_2(p_x, p_y, p_z), \\ \frac{p_z C + (p_x + ip_y)D}{W/c + m_0 c} + A = g_3(p_x, p_y, p_z), \quad \frac{p_z D - (p_x - ip_y)C}{W/c + m_0 c} + B = g_4(p_x, p_y, p_z). \end{aligned}$$

If one studies the preceding system, while taking the Heisenberg uncertainty relations into account, then one will see that in the case of a wave packet that is either immobile or animated with a velocity $v = c$, one will get the following result: It is possible to represent the wave packet ψ by a superposition of (positive-energy) monochromatic plane waves if and only if that wave packet has dimensions that are equal to at least $h / m_0 c$. If the wave packet has dimensions that are less than $h / m_0 c$ then the superposition must include negative-energy states. If the wave packet has a velocity βc that is close to c in the reference system that is being used then upon taking the Lorentz contraction into account one will easily see that the wave packet will be represented by a superposition of positive-energy monochromatic plane waves iff its dimensions have order at least $\frac{h}{m_0 c} \sqrt{1 - \beta^2}$.

In general, it is therefore impossible to represent an arbitrary wave train in Dirac's theory without appealing to negative-energy waves, which shows that one cannot avoid those waves.

The condition for one to be able to represent a wave train uniquely by positive-energy waves is then $\delta x \gg \frac{h}{m_0 c} \sqrt{1 - \beta^2}$, in which δx is the extension of the wave train along an arbitrary axis Ox . We shall recover that condition by starting from the uncertainty relations. The fourth uncertainty relation:

$$\delta W \cdot \delta t > h$$

must be interpreted by saying that if a measurement of the energy provides the value of energy, up to an error of δW , then a measurement of the time duration δt must equal at least $h / \delta W$. In order for the wave train considered to contain only positive-energy waves in its Fourier development, it will be necessary that $\delta W < W$, since from the principles of wave mechanics, δW must measure the extent (at the scale of W) of the waves that appear in the Fourier development of ψ . The duration δt of the measurement that permits one to delimit the wave train will therefore be $\delta t > h / W$. Now, in Dirac's wave mechanics, the wave fronts can propagate with any velocity from 0 to c , in such a way that during the duration δt of the measurement, the boundary of the wave packet can displace by $c \delta t$. The extent of the wave packet that results from the measurement will be at least $\delta x = c \delta t$ in the Ox direction. One will then have:

$$\delta x > \frac{hc}{W} = \frac{h}{m_0 c} \sqrt{1 - \beta^2},$$

since:

$$W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}}.$$

That is, in fact, the result that was obtained above.

If one applies the formula $\delta t > hc / W$ to light (for which $W = h\nu$ and $\lambda = c / \nu$) then one will find that $\delta\nu > \lambda$. It is impossible to localize a photon within a volume whose dimensions are smaller than the wave length. On the contrary, for a material particle (the electron, for example), one will have $\delta x > \frac{h}{m_0\nu} \sqrt{1 - \beta^2} \frac{\nu}{c} = \lambda \cdot \beta$, and since $\beta < 1$, one can localize the particle in a volume whose dimensions are smaller than the associated wave length, and that will be all the more true as the velocity becomes smaller.

3. DIRAC'S THEORY OF "HOLES"

From our study of the Klein paradox and the incomplete character of the positive-energy plane waves, we just saw that it is impossible to eliminate the negative-energy plane waves from Dirac's theory. We further note the following curious result: The Dirac electron cannot diffuse light unless the negative-energy states exist. The existence of the phenomenon of the diffusion of light will then also prevent us from eliminating those paradoxical waves.

Dirac himself has found an ingenious means of eliminating that difficulty, and in his theory, that consists of the existence of negative-energy waves that are impossible to discard. We remark that from the Pauli Exclusion Principle, one cannot have more than one electron in a state, so he imagined that all of the negative-energy states are filled in the normal state of the universe. A uniform density of negative-energy electrons in the universe will then result, and Dirac assumed that the normal uniform density was unobservable. However, there are more electrons in the universe than is necessary to fill all of the negative-energy states, and the surplus, which one is obliged to distribute between the positive-energy states, will constitute the electrons that manifest themselves in experiments. In the exceptional case, a negative-energy electron can pass to a positive-energy state under the action of an external force. An experimental electron and a "hole" will then appear simultaneously, which is a gap in the distribution of negative-energy electrons. Now, Dirac showed (and we shall return to this point later on) that such a gap will behave like a corpuscle with positive energy that has the same mass as the electron and an electric charge that is equal, but of opposite sign. It will be a corpuscle that is, in some sense, *complementary* to the usual electron, namely, the positive electron or "positon."

In 1932, Anderson, along with Blackett and Occhialini, discovered positive electrons in cosmic rays that answered to Dirac's conception of them. One will also discover positons in the products of the disintegration of artificial radio-elements. Despite the rather exceptional character of their appearance, positons have become quite familiar to physicists today. Dirac's theory of holes leads one to think that the positive electrons must be unstable in the presence of matter, because if a gap meets a negative-energy electron in matter where it is full then that negative electron can combine with the gap, and the quantum transition will translate into the simultaneous disappearance of two electrons of opposite signs whose total energy will be emitted in the form of radiation. That is the phenomenon of "the dematerialization of a pair of electrons" whose existence is certain today, thanks to, above all, the work of Joliet and Thibaud.

With Dirac, we can then consider the positon to be a “hole” or “gap” in the distribution of the negative-energy state of the electron, but, more physically, we can also consider it to be a true corpuscle that is “complementary” to the electron, and whose wave function φ will obey a wave equation that is complementary to that of Dirac. Let us make that point more precise.

Consider the case of the absence of a field, and let ψ^+ be a positive-energy solution of the Dirac equations:

$$\frac{1}{c} \frac{\partial \psi_k^+}{\partial t} = \left(\alpha_1 \frac{\partial}{\partial x} + \alpha_2 \frac{\partial}{\partial y} + \alpha_3 \frac{\partial}{\partial z} + \frac{2\pi i}{h} m_0 c \alpha_4 \right) \psi_k^+.$$

I say that the function $\psi^- = -i \alpha_2 \alpha_4 (\psi^+)^*$, whose components are:

$$\psi_1^- = -(\psi_1^+)^*, \quad \psi_2^- = (\psi_3^+)^*, \quad \psi_3^- = (\psi_2^+)^*, \quad \psi_4^- = -(\psi_4^+)^*,$$

because:

$$-i \alpha_2 \alpha_4 = \begin{vmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{vmatrix}$$

is a *negative-energy* solution of the Dirac equations. The fact that it has negative energy is obvious. One sees that it is a solution of the Dirac equations by writing the conjugate equation:

$$\frac{1}{c} \frac{\partial \psi_k^{+*}}{\partial t} = \left(\alpha_1 \frac{\partial}{\partial x} - \alpha_2 \frac{\partial}{\partial y} + \alpha_3 \frac{\partial}{\partial z} - \frac{2\pi i}{h} m_0 c \alpha_4 \right) \psi_k^{+*}$$

that is obtained by noting that with our choice of α_i , α_2 , α_3 , and α_4 are real, while α_2 is pure imaginary, and then left-multiplying the preceding equation by $-i \alpha_2 \alpha_4$ and taking into account the commutation relations between the α_i .

Hence, if we start with a positive-energy solution ψ^+ then we can *associate* it with a negative-energy solution ψ^- . Now set:

$$\varphi = \psi^{-*} = -i \alpha_2 \alpha_4 \psi^+,$$

with the components:

$$\varphi_1 = \psi_1^+, \quad \varphi_2 = \psi_2^+, \quad \varphi_3 = \psi_3^+, \quad \varphi_4 = \psi_4^+.$$

When one left-multiplies the equation for ψ^+ by $-i \alpha_2 \alpha_4$, it is easy to verify that φ will be a solution to:

$$\frac{1}{c} \frac{\partial \varphi_k}{\partial t} = \left(\alpha_1 \frac{\partial}{\partial x} - \alpha_2 \frac{\partial}{\partial y} + \alpha_3 \frac{\partial}{\partial z} - \frac{2\pi i}{h} m_0 c \alpha_4 \right) \varphi_k,$$

which we call the *equation that is complementary to the Dirac equation*. The function φ will be the wave function for a corpuscle that is “complementary” to the electron; i.e., the positon. It is a solution to the “complementary” Dirac equations.

We now pass on to the case in which an electromagnetic field exists that is derivable from the potentials V and \mathbf{A} . Once more, let ψ^+ be a *positive-energy* solution to the Dirac equations such that:

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \kappa e \frac{V}{c}\right)\psi_k^{+*} = \left[\alpha_1\left(\frac{\partial}{\partial x} + \kappa e A_x\right) + \alpha_2\left(\frac{\partial}{\partial y} + \kappa e A_y\right) + \alpha_3\left(\frac{\partial}{\partial z} + \kappa e A_z\right) + \kappa m_0 c \alpha_4\right]\psi_k^{+*},$$

with:

$$\kappa = \frac{2\pi i}{h}.$$

In order to obtain the negative-energy solution ψ^- that is associated with ψ^+ , we begin by considering the equation that is obtained by starting with the preceding one and replacing e with $-e$. We will get a solution ψ'^+ to it by changing e into $-e$ in the expression for ψ^+ , and we will have:

$$\left(\frac{1}{c}\frac{\partial}{\partial t} - \kappa e \frac{V}{c}\right)\psi_k'^+ = \left[\alpha_1\left(\frac{\partial}{\partial x} - \kappa e A_x\right) + \alpha_2\left(\frac{\partial}{\partial y} - \kappa e A_y\right) + \alpha_3\left(\frac{\partial}{\partial z} - \kappa e A_z\right) + \kappa m_0 c \alpha_4\right]\psi_k'^+.$$

The desired solution ψ^- will then be:

$$\psi^- = -i \alpha_2 \alpha_4 (\psi'^+)^*.$$

One will easily see this by taking the conjugate of the last equation:

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \kappa e \frac{V}{c}\right)(\psi_k'^+)^* = \left[\alpha_1\left(\frac{\partial}{\partial x} + \kappa e A_x\right) + \alpha_2\left(\frac{\partial}{\partial y} + \kappa e A_y\right) + \alpha_3\left(\frac{\partial}{\partial z} + \kappa e A_z\right) - \kappa m_0 c \alpha_4\right](\psi_k'^+)^*,$$

and applying the operator $-i \alpha_2 \alpha_4$ on the left, which will give:

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \kappa e \frac{V}{c}\right)\psi_k^- = \left[\alpha_1\left(\frac{\partial}{\partial x} + \kappa e A_x\right) + \alpha_2\left(\frac{\partial}{\partial y} + \kappa e A_y\right) + \alpha_3\left(\frac{\partial}{\partial z} + \kappa e A_z\right) + \kappa m_0 c \alpha_4\right]\psi_k^-.$$

Naturally, if:

$$V = 0, \quad \mathbf{A} = 0 \quad \text{then:} \quad \psi'^+ = \psi^+ \quad \text{and} \quad \psi^- = -i \alpha_2 \alpha_4 \psi^{+*},$$

in such a way that we will recover the result that was established above.

Set:

$$\varphi = \psi^{-*} = -i \alpha_2 \alpha_4 \psi'^+.$$

Upon taking the conjugate of the equation that is satisfied by ψ^- , we will find that:

$$\left(\frac{1}{c} \frac{\partial}{\partial t} - \kappa e \frac{V}{c}\right) \varphi_k = \left[\alpha_1 \left(\frac{\partial}{\partial x} - \kappa e A_x \right) + \alpha_2 \left(\frac{\partial}{\partial y} - \kappa e A_y \right) + \alpha_3 \left(\frac{\partial}{\partial z} - \kappa e A_z \right) + \kappa m_0 c \alpha_4 \right] \varphi_k .$$

The wave φ represents a corpuscle that obeys the complementary Dirac equation and *possesses the charge + e*. It represents a positron, which is the corpuscle that is complementary to the electron.

Naturally, for $V = 0$, $\mathbf{A} = 0$, we will get back to the results that were obtained above.

CHAPTER VIII

SCHRÖDINGER'S ZITTERBEWEGUNG

1. THE CENTER OF GRAVITY OF PROBABILITY IN WAVE MECHANICS

In wave mechanics, one must avoid speaking of velocity, more generally, because velocity is defined only in special cases (e.g., monochromatic plane waves). Instead of velocity, one must speak of the quantity of motion. By contrast, one must always define a point G that is called the *center of gravity of the probability of presence*, and which is defined by the formulas:

$$\bar{x} = \int_D \psi^* x \psi d\tau, \quad \bar{y} = \int_D \psi^* y \psi d\tau, \quad \bar{z} = \int_D \psi^* z \psi d\tau$$

in non-relativistic wave mechanics, and by the formulas:

$$\bar{x} = \int_D \sum_{k=1}^4 \psi_k^* x \psi_k d\tau, \quad \bar{y} = \int_D \sum_{k=1}^4 \psi_k^* y \psi_k d\tau, \quad \bar{z} = \int_D \sum_{k=1}^4 \psi_k^* z \psi_k d\tau$$

in Dirac's wave mechanics; $\bar{x}, \bar{y}, \bar{z}$ are, by definition, the coordinates of that point.

In the course of time, $\bar{x}, \bar{y}, \bar{z}$ are well-defined functions of time, in such a way that one can unambiguously speak of the velocity of the point G .

Let $\psi_1, \psi_2, \psi_3, \psi_4$ be a solution be a solution of the Dirac equations that represents the wave that is associated with a certain motion of the electron. We can write:

$$\psi_k = \sum c_n \psi_{k,n},$$

in which $\psi_{k,n}$ is the k^{th} component of the n^{th} proper function of energy (viz., a proper function of H), and the c_n are complex constants. One finds, by substitution, that:

$$\bar{x} = \sum_{m,n} c_m^* c_n \int_D \sum_{k=1}^4 \psi_{k,m}^* x \psi_{k,n} d\tau = \sum_{m,n} c_m^* c_n x_{mn},$$

in which $x_{m,n}$ is the element of the Heisenberg matrix that corresponds to the operator x .

Now:

$$\frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^4 \left[\psi_{k,m}^* \frac{2\pi i}{h} (xH - Hx) \psi_{k,n} \right] d\tau,$$

with

$$H = - [eV + c (\alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 P_4)].$$

H is the Dirac Hamiltonian in the presence of a field, as one knows:

$$P_j = -\frac{h}{2\pi i} \frac{\partial}{\partial x_j} + \frac{e}{c} A_j.$$

One easily finds that:

$$\frac{2\pi i}{h} (xH - Hx) = c\alpha_1 \left(x \frac{\partial}{\partial x} - \frac{\partial}{\partial x} x \right) = -c\alpha_1,$$

and as a result:

$$\frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^4 \left[\psi_{k,m}^* \frac{2\pi i}{h} (-c\alpha_1) \psi_{k,n} \right] d\tau,$$

so:

$$\frac{d\bar{x}}{dt} = \sum_{m,n} c_m^* c_n \frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^4 \left[\sum_m c_m^* \psi_{k,m}^* (-c\alpha_1) \sum_n c_n \psi_{k,n} \right] d\tau = \int_D \sum_{k=1}^4 \left[\psi_k^* (-c\alpha_1) \psi_k \right] d\tau,$$

and one will likewise find that:

$$\frac{d\bar{y}}{dt} = \int_D \sum_{k=1}^4 \left[\psi_k^* (-c\alpha_2) \psi_k \right] d\tau, \quad \frac{d\bar{z}}{dt} = \int_D \sum_{k=1}^4 \left[\psi_k^* (-c\alpha_3) \psi_k \right] d\tau.$$

Now, we have previously found that the expressions for the components of the probability current are:

$$\rho u_x = -c \sum_{k=1}^4 \psi_k^* \alpha_1 \psi_k, \dots,$$

so:

$$\frac{d\bar{x}}{dt} = \int_D \rho u_x d\tau = \bar{u}_x, \quad \frac{d\bar{y}}{dt} = \bar{u}_y, \quad \frac{d\bar{z}}{dt} = \bar{u}_z.$$

The velocity of the center of gravity G is then equal to the mean value of the velocity of the probability, which is a result that one could predict *a priori*.

One often interprets the preceding formulas by saying that the operators $-c\alpha_1$, $-c\alpha_2$, and $-c\alpha_3$ are the operators that correspond to the three components of the velocity of the corpuscle. Since those operators have only $+c$ and $-c$ for proper values, one will be led to say that from the general principles of wave mechanics, the only possible values for the components of velocity will be $+c$ and $-c$, which is a result that is very difficult to interpret. However, it seems preferable to not consider the velocity to be an “observable” to which one can associate an operator.

2. EHRENFEST'S THEOREM

In non-relativistic wave mechanics, one proves Ehrenfest's theorem, which is expressed by the formulas:

$$m \frac{d^2 \bar{x}}{dt^2} = \bar{f}_x, \quad m \frac{d^2 \bar{y}}{dt^2} = \bar{f}_y, \quad m \frac{d^2 \bar{z}}{dt^2} = \bar{f}_z,$$

which are made more specific by saying that the center of gravity of the probability displaces like a material point of mass m under the influence of the mean value of the force:

$$\bar{\mathbf{f}} = \int_D \psi^* (-\text{grad } U) \psi d\tau.$$

In the case where there is no field, the point G will then be animated with a uniform, rectilinear motion, so it will not be at rest. We shall prove the latter theorem directly. In

non-relativistic wave mechanics $H = \left(\frac{h}{2\pi i}\right)^2 \frac{1}{2m} \Delta$, so:

$$\frac{d\bar{x}}{dt} = \frac{d}{dt} \int_D \psi^* x \psi d\tau = \int_D \psi^* \frac{2\pi i}{h} [xH - Hx] \psi d\tau = \int_D \psi^* \left(\frac{-h}{2\pi i} \frac{\partial}{\partial x}\right) \psi d\tau,$$

because:

$$[x, \Delta] = -2 \frac{\partial}{\partial x},$$

and as a result:

$$\frac{d^2 \bar{x}}{dt^2} = \int_D \psi^* \frac{h}{2\pi i} \left[-\frac{h}{2\pi i} \frac{\partial}{\partial x} H - H \left(\frac{-h}{2\pi i}\right) \frac{\partial}{\partial x} \right] \psi d\tau = 0,$$

since $\partial / \partial x$ commutes with Δ . In Dirac's theory, the preceding results will no longer be exact, in general.

One infers from:

$$\frac{d\bar{x}}{dt} = \int_D \sum_{k=1}^4 [\psi_k^* (-c\alpha_1) \psi_k] d\tau$$

that:

$$\frac{d^2 \bar{x}_{mn}}{dt^2} = \int_D \sum_{k=1}^4 \psi_{k,m}^* \frac{h}{2\pi i} [-c\alpha_1 H + H c\alpha_1] \psi_{k,m} d\tau.$$

However, α_1 does not commute with H here, since it contains α_2 , α_3 , and α_4 . One will then have:

$$\frac{d^2 \bar{x}}{dt^2} \neq 0, \quad \text{and also} \quad \frac{d^2 \bar{y}}{dt^2} \neq 0 \quad \text{and} \quad \frac{d^2 \bar{z}}{dt^2} \neq 0.$$

The motion of the center of gravity of the probability in the absence of field is not generally uniform and rectilinear.

3. SCHRÖDINGER'S ZITTERBEWEGUNG

In order to understand why the motion of the center of gravity of probability in Dirac's theory is not uniform and rectilinear, even in the absence of a field, it is instructive to subject the expression:

$$\frac{d\bar{x}}{dt} = \int_D \sum_{k=1}^4 [\psi_k^* (-c\alpha_1) \psi_k] d\tau$$

to a detailed analysis, in which we can write:

$$\psi_k = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} + b_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)} \right] \cdot dp_x dp_y dp_z,$$

with $W = +c \sqrt{m_0^2 c^2 + p^2}$. The eight a_k and b_k can be calculated by starting with four of them, and they will remain arbitrary from the formulas that we know.

Now, say *momentum space* to mean the space that is constructed by taking the quantities p_x, p_y, p_z to be its rectangular coordinates, and divide that space into cells σ that are as small as one desires. The quantities:

$$\Delta(\sigma) = \iiint e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dp_x dp_y dp_z$$

are (up to a normalization constant) the "proper differentials" of the continuous spectrum of the monochromatic plane waves, and we can write:

$$\psi_k(x, y, z, t) = \sum_{\sigma} \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h} Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h} Wt} \right] \Delta(\sigma),$$

in which p_x, p_y, p_z are the coordinates of the center of the element σ in momentum space, and \sum_{σ} denotes a summation over all of the cells σ in that space.

We can then write:

$$\begin{aligned} \frac{d\bar{x}}{dt} = & -c \sum_{\sigma} \sum_{\sigma'} \sum_{k=1}^4 \left[a_k(p'_x, p'_y, p'_z) e^{\frac{2\pi i}{h} Wt} + b_k(p'_x, p'_y, p'_z) e^{-\frac{2\pi i}{h} Wt} \right]^* \\ & \cdot \alpha_1 \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h} Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h} Wt} \right] \int_D \Delta^*(\sigma') \Delta(\sigma) d\tau, \end{aligned}$$

in which the domain D is naturally the entire space, here. Since the proper differentials are orthogonal and assumed to be normalized, we will have, upon letting the volume of the cell σ be σ :

$$\frac{d\bar{x}}{dt} = -c \sum_{\sigma} \sigma \sum_{k=1}^4 \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i W t}{h}} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i W t}{h}} \right]^* \\ \cdot \alpha_1 \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i W t}{h}} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i W t}{h}} \right],$$

or furthermore:

$$\frac{d\bar{x}}{dt} = -c \sum_{\sigma} \sigma \sum_{k=1}^4 (a_k^* \alpha_1 a_k + b_k^* \alpha_1 b_k) - c \sum_{\sigma} \sigma \left(\sum_{k=1}^4 a_k^* \alpha_1 b_k e^{\frac{2\pi i W t}{h}} + \sum_{k=1}^4 b_k^* \alpha_1 a_k e^{-\frac{2\pi i W t}{h}} \right).$$

Upon taking formulas (VII.b) into account, which we recall here:

$$a_1 = -\frac{p_z A + (p_x + ip_y) B}{W/c + m_0 c}, \quad a_2 = -\frac{p_z B - (p_x - ip_y) A}{W/c + m_0 c}, \quad a_3 = A, \quad a_4 = B,$$

we will have:

$$-c \sum_{k=1}^4 a_k^* \alpha_1 a_k = -c (a_1^* a_4 + a_2^* a_3 + a_3^* a_2 + a_4^* a_1) = 2p_x c \frac{AA^* + BB^*}{W/c + m_0 c},$$

and since we also have:

$$\sum_{k=1}^4 a_k^* a_k = (A^* A + B^* B) \left[1 + \frac{p^2}{(W/c + m_0 c)^2} \right] = 2W \frac{AA^* + BB^*}{W/c + m_0 c},$$

we will get:

$$-c \sum_{k=1}^4 a_k^* \alpha_1 a_k = \frac{p_x c^2}{W} \sum_{k=1}^4 a_k^* a_k.$$

On the other hand, the last two terms in the expression for $d\bar{x}/dt$ are complex-conjugates, due to the Hermiticity of α_1 , and can be written:

$$c \sum_{\sigma} \sigma A_1 \cos\left(\frac{4\pi}{h} W t + \varphi_1\right).$$

Naturally, A_1 and φ_1 vary from one cell σ to the other – i.e., they are functions of p_x, p_y, p_z . Indeed, one can set:

$$A_1 e^{i\varphi_1} = -2 \sum_{k=1}^4 b_k^* \alpha_1 a_k.$$

Finally, we can write:

$$\frac{d\bar{x}}{dt} = \sum_{\sigma} \sigma \frac{c^2 p_x}{W} \left(\sum_{k=1}^4 a_k^* a_k - \sum_{k=1}^4 b_k^* b_k \right) + \sum_{\sigma} \sigma c A_1 \cos\left(\frac{4\pi}{h} W t + \varphi_1\right).$$

Now, from the formulas of relativistic dynamics, $p_x c^2 / W$ is the x -component of the velocity that corresponds to the quantity of motion p_x and the energy $+W$. One can similarly consider $-p_x c^2 / W$ to be the x -component of the velocity that corresponds to the quantity of motion p_x and energy $-W$. The first term in the preceding expression for $d\bar{x}/dt$ is then a sort of mean value of the component of the velocity v_x that corresponds to the spectral decomposition of ψ .

Therefore, set:

$$\bar{v}_x = \sum_{\sigma} \sigma \frac{c^2 p_x}{W} \sum_{k=1}^4 (a_k^* a_k - b_k^* b_k).$$

\bar{v}_x is naturally independent of time, and we will have:

$$\frac{d\bar{x}}{dt} = \bar{v}_x + \sum_{\sigma} \sigma c A_1 \cos\left(\frac{4\pi}{h} Wt + \varphi_1\right).$$

Hence, upon integration:

$$\bar{x} = \xi + \sum_{\sigma} \sigma \frac{hc}{4\pi W} A_1 \sin\left(\frac{4\pi}{h} Wt + \varphi_1\right),$$

and one will likewise find that:

$$\begin{aligned} \bar{y} &= \eta + \sum_{\sigma} \sigma \frac{hc}{4\pi W} A_2 \sin\left(\frac{4\pi}{h} Wt + \varphi_1\right), \\ \bar{z} &= \zeta + \sum_{\sigma} \sigma \frac{hc}{4\pi W} A_3 \sin\left(\frac{4\pi}{h} Wt + \varphi_1\right), \end{aligned}$$

with the definitions:

$$\begin{aligned} \xi &= \bar{x}_0 + \bar{v}_x t, & \eta &= \bar{y}_0 + \bar{v}_y t, & \zeta &= \bar{z}_0 + \bar{v}_z t, \\ \bar{v}_y &= \sum_{\sigma} \sigma \frac{c^2 p_y}{W} \sum_{k=1}^4 (a_k^* a_k - b_k^* b_k), & \bar{v}_z &= \sum_{\sigma} \sigma \frac{c^2 p_z}{W} \sum_{k=1}^4 (a_k^* a_k - b_k^* b_k). \end{aligned}$$

The point whose coordinates are ξ , η , ζ then displaces with a uniform, rectilinear motion, but the center of gravity G , whose coordinates are \bar{x} , \bar{y} , \bar{z} , will execute a series of oscillations around that point whose frequency is $2W/h$. That is *Schrödinger's Zitterbewegung*, which prevents Ehrenfest's theorem from being valid. Moreover, the amplitudes of those oscillations are weak, in general, because they are proportional to the factor $hc/4\pi W$, which is always smaller than $\frac{h}{4\pi m_0 c} = \frac{1}{4\pi} \frac{h}{m_0 c}$. Now, the quantity

$\frac{h}{m_0 c}$ is often called the *Compton wavelength*, and is very small (e.g., 2.4×10^{-10} cm for the electron).

The preceding analysis shows neatly the origin of Schrödinger's *Zitterbewegung*: It is due to the beating of waves of positive energy $+W$ with waves of negative energy $-W$. As usual, the beat frequency is the difference of the frequencies, namely, $2W/h$.

For a wave train in a special decomposition for which no negative-energy waves appear, there will be no Schrödinger *Zitterbewegung*, because if $b_k = 0$ then $A_1 = 0$, and as a result, Ehrenfest's theorem will remain valid. However, we know that in order to represent a wave train, one must generally introduce negative-energy waves (if the dimensions of the train are less than $\frac{h}{m_0c} \sqrt{1-\beta^2}$), and that is why Ehrenfest's theory will not be valid in Dirac's theory.

CHAPTER IX

POSSIBILITY OF MEASURING THE SPIN OF THE ELECTRON

1. PRESENT THINKING ABOUT THE QUESTION

Bohr gave some arguments to show that it is impossible to exhibit the spin of the electron *by direct measurements*. Naturally, that will not exclude the *indirect* measurement of that spin by confirming its repercussions on various phenomena, such as the fine structure of spectra. It no longer excludes the possibility of doing something that does not seem to have been done in a clean fashion up to the present, namely, exhibiting the state of polarization of an electron wave by experiments of the Norremberg type in optics (viz., reflecting a wave that was already polarized by a previous reflection from a polarizing body).

We shall study Bohr's argument (which was developed notably by Pauli in the Actes de Congrès Solvay, 1930, pp. 217, *et seq.*) and see that the validity seems to be generally limited to the case of velocities that are small compared to c .

2. ACTION OF A MAGNETIC FIELD UPON THE PROPER MAGNETIC MOMENT

Consider a parallel, monokinetic, sheaf of particles whose velocity is directed along Ox and whose dimensions transverse to the sheaf are Δy and Δz . One can suppose, for example, that the sheaf has been bounded laterally by passing it through a rectangular slit with edges Δy and Δz . We suppose that the particles, which are not electrons, although they are particles of spin 1/2 that obey the Dirac equations, have a charge of ε and a proper mass m_0 , and that they are subject to the action of an inhomogeneous magnetic field that is everywhere parallel to the plane yOz , in such a way that we will have:

$$\operatorname{div} \mathbf{H} = \frac{\partial H_y}{\partial y} + \frac{\partial H_z}{\partial z} = 0.$$

If we fix our attention upon the actions that are exerted upon the particles along Oz (an analogous argument applies to Oy) then we will see that the Laplacian [*sic*] action of the magnetic field upon the particle translates into the existence of a force $(e/c) v H_y$ along Oz , while the action of the proper magnetic moment \mathfrak{M} of the particle will give rise to the force $\mathfrak{M}_z \partial H_z / \partial z$, and when one takes into account the fact that $\operatorname{div} \mathbf{H} = 0$, it will be a force that has the maximum absolute value $\frac{\varepsilon h}{4\pi m_0 c} \cdot \frac{\partial H_y}{\partial y}$ in the case where the proper magnetic moment is normal to the axis and directed along Oz .

However, the sheaf has a size Δy in the direction Oy , so the Laplacian [*sic*] force will be affected with an uncertainty that is equal to $\frac{\epsilon}{c} v \cdot \frac{\partial H_y}{\partial y} \Delta y$, and in order for one to be able to exhibit the existence of proper magnetic moment, it is necessary that the uncertainty should be much smaller than the force that is due to the gradient of the field, which will lead to the inequality:

$$\Delta y \ll \frac{1}{4\pi} \frac{h}{m_0 v}.$$

For particles whose velocities are small compared to c , the factor $h / m_0 v$ will be equal to the wave length λ of the associated wave, and one will have:

$$\Delta y \ll \lambda.$$

That condition implies the existence of an intense diffraction that will no longer allow one to attribute a well-defined trajectory to the particle, which will result in the impossibility of exhibiting the existence of the proper magnetic moment in an experiment for which the notion of trajectory must be preserved.

It is nonetheless convenient to remark that in the preceding argument, we have supposed that the particle of mass m_0 carries a magnetic moment $\frac{\epsilon h}{4\pi m_0 c}$ that

corresponds to it in Dirac's theory. Things are different in the case of the Stern-Gerlach experiment, in which a silver atom of total proper mass M_0 carries a magnetic moment that equal to the Bohr magneton $\frac{\epsilon h}{4\pi m_0 c}$, where m_0 is the mass of the electron in that

case. Upon repeating the previous argument and remarking that $\lambda = h / M_0 v$, one will find that:

$$\Delta y \ll \frac{M_0}{m_0} \lambda,$$

in place of:

$$\Delta y \ll \lambda,$$

and since m_0 / M_0 is very small, nothing will stand in the way of our being able to exhibit the proper magnetic moment of the electron in experiments where one can attribute a trajectory to a silver atom.

We now come to an essential point. We previously neglected the relativistic corrections by setting $\lambda = h / M_0 v$, whereas the rigorous expression for the wave length is:

$$\lambda = \frac{h}{m_0 v} \sqrt{1 - \beta^2} \quad (\beta = v / c).$$

If we take that rigorous expression into account then we will get:

$$\Delta y \ll \frac{1}{4\pi} \cdot \frac{\lambda}{\sqrt{1-\beta^2}}$$

instead of $\Delta y \ll \lambda$, and if β is sufficiently close to 1 (i.e., v is sufficiently close to c) then $\sqrt{1-\beta^2}$ will be very small, and the preceding condition will no longer imply that Δy is very small with respect to λ , which is a condition that was proved above in the case of $v \ll c$.

One cannot then assert that it will be impossible to exhibit the proper magnetic moment of Dirac particles that are animated with a velocity that is close to c by the experiment that is envisioned.

Meanwhile, as Thibaud has remarked, a *practical* impossibility persists for electrons, even when $v \sim c$. Indeed, one has $\Delta y = h / m_0 v$, which will give $\Delta y \ll 10^{-10}$ cm for electrons, and one cannot use a slit that is narrow enough. Things will be different for particles that are much lighter than electrons. For example, for $m \sim 10^{-37}$ g = $10^{-10} m_0$, one will have $\Delta y \ll 1$ cm, which is easy to realize.

In summation, for slow particles ($v \ll c$), it might seem impossible to exhibit the proper magnetic moment. For rapid particles ($v \sim c$), there will no longer be any theoretical impossibility (from the uncertainty relations), but there will always be a practical impossibility (from the experimental arrangements that one can utilize). For particles that are much lighter than electrons with velocities that are close to c , one will no longer perceive *a priori* that impossibility of exhibiting the proper magnetic moment.

Note: The component of magnetic moment that is normal to the velocity is independent of the velocity because it is equal to:

$$\int \mu_{23} d\tau = \int \frac{\mu_{23}^0}{\sqrt{1-\beta^2}} d\tau_0 \sqrt{1-\beta^2} = \int \mu_{23}^0 d\tau.$$

3. MEASURING THE MAGNETIC FIELD THAT IS PRODUCED BY THE ELECTRON

The magnetic field that is produced by the magnetic electron at a point at a distance r has order of magnitude:

$$H_1 = \frac{e}{r^2} \cdot \frac{v}{c}, \quad H_2 = \frac{\mathfrak{M}}{r^3},$$

with:

$$\mathfrak{M} = \text{proper magnetic moment} = \frac{\varepsilon h}{4\pi m_0 c}, \quad \varepsilon = -e.$$

For an uncertainty Δv in v , one will have:

$$\Delta H_1 = \frac{e}{cr^2} \Delta v.$$

Now, the position of the electron is uncertain to Δr , so one will have:

$$\Delta p \Delta r \gg h,$$

because Δr can be considered to be the size of a slit that the electron crosses normally. Moreover, any precise measurement will assume that $\Delta r \ll r$. If one assumes the *non-relativistic* formula $p = m_0 v$ then one will have:

$$\Delta v \gg \frac{h}{m_0 \Delta r} \quad \text{and} \quad \Delta H_1 \gg \frac{e}{cr^2} \cdot \frac{h}{m_0 \Delta c} \gg \frac{eh}{m_0 cr^3} \quad \text{or} \quad \Delta H_1 \gg \frac{\mathfrak{M}}{r^3}.$$

The uncertainty in H_1 is then much higher than H_2 , which will prevent one from measuring H_2 , but here again the same thing will not be true when $v \sim c$, because then one will have $p = \frac{m_0 v}{\sqrt{1-\beta^2}}$, and:

$$\Delta p = \frac{m_0 v}{\sqrt{1-\beta^2}} \Delta v + m_0 v \Delta \left(\frac{1}{\sqrt{1-\beta^2}} \right) = m_0 \Delta v \left[\frac{1}{\sqrt{1-\beta^2}} + \frac{v}{\Delta v} \Delta \left(\frac{1}{\sqrt{1-\beta^2}} \right) \right] \gg m_0 \Delta v,$$

because:

$$\frac{v}{\Delta v} > 1, \quad \frac{1}{\sqrt{1-\beta^2}} \gg 1, \quad \Delta \left(\frac{1}{\sqrt{1-\beta^2}} \right) \cdot \frac{v}{\Delta v} = \frac{v^2}{c^2} \cdot \frac{1}{(1-\beta^2)^{3/2}} \gg 1.$$

Set $\Delta p = m_0 \Delta v \cdot N$, with $N \gg 1$, so one will get:

$$\Delta H_1 \gg \frac{e}{cr^2} \cdot \frac{h}{m_0 N \Delta r} \gg \frac{eh}{m_0 cr^3} \cdot \frac{1}{N},$$

but since N is very large, *we can no longer deduce that:*

$$\Delta H_1 \gg H_2.$$

4. COMPENSATING FOR THE LORENTZ FORCE WITH AN ELECTRIC FIELD

Pauli has also imagined the following case: Some electrons are emitted with velocity v along the z -axis, and are subject to the action of an electric field and a magnetic field, both of which are parallel to the plane xOy and independent of z . Those fields are derived

from a scalar potential V and a vector potential A_z that are functions of only x and y , and are such that:

$$E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}, \quad H_x = \frac{\partial A_z}{\partial x}, \quad H_y = \frac{\partial A_z}{\partial y}.$$

The force that acts upon the electrons has the components:

$$F_x = -e E_x + \frac{e}{c} v H_y = e \frac{\partial}{\partial x} \left(V - \frac{v}{c} A_z \right),$$

$$F_y = -e E_y - \frac{e}{c} v H_x = e \frac{\partial}{\partial y} \left(V - \frac{v}{c} A_z \right).$$

The forces will be compensated *in all space* if $V = v/c A_z$. Pauli said that it would then seem that one could exhibit the forces that are exerted upon the proper magnetic moment, which would be the only ones to persist, while an exact definition of the position of the electrons in the directions that are perpendicular to the sheaf would not be necessary. However, if $v_x = v_y = 0$, v_z cannot be independent of x and y , because it is not mv_z , but in fact $p_z = mv_z - (e/c) A_z$, which is a first integral in relativistic dynamics (m is the *moving* mass $m_0 / \sqrt{1 - \beta^2}$).

Suppose that p_z is constant, and imagine a sheaf of electrons that is bounded laterally by a circular opening of diameter d . Furthermore, let:

$$A_z = -(ax + byx), \quad H_x = -bx, \quad H_y = a + by, \quad \frac{\partial H_y}{\partial y} = b.$$

Along the z -axis, the magnetic field:

$$H = H_y = a$$

will orient the proper magnetic moments, and the force that is exerted upon each of those moments will be $F_y = \mathfrak{M} b$. On the boundary of the sheaf, there is an *uncompensated* Lorentz force:

$$\Delta F_y = \frac{e}{c} \Delta v_z H_x = \frac{e}{c} \Delta v_z b d,$$

and since:

$$\Delta p_z = 0 \quad \text{or} \quad \Delta v_z = \frac{e}{mc} \Delta A = \frac{e}{mc} b d^2,$$

one will have:

$$\Delta F_y = \frac{e}{mc} \cdot \frac{e}{c} b^2 d^2,$$

in which m is the mass of motion $m_0 / \sqrt{1-\beta^2}$. In order to be able to measure F_y , it is necessary that $\Delta F_y \ll F_y$, or:

$$\mathfrak{M} b \gg \frac{e}{mc} \cdot \frac{e}{c} b^2 d^2,$$

or rather:

$$\mathfrak{M} \gg \frac{e}{mc} \cdot \frac{e}{c} b d^2.$$

The angle of diffraction of the sheaf after passing through a hole of diameter d has order λ / d ($\lambda = h / mv_z$), and in order for the notion of trajectory to preserve any sense, it is necessary that the distance l that is travelled in the field must be such that:

$$\frac{\lambda}{d} l \ll d \quad \text{or} \quad 1 \ll \frac{\lambda l}{d^2}.$$

Upon multiplying that inequality by the preceding one, one will find that:

$$\mathfrak{M} \gg \frac{e}{mc} \cdot \frac{e}{c} \cdot b d \lambda l$$

or:

$$\frac{\lambda l}{d} \mathfrak{M} \gg \frac{e}{mc} \cdot \frac{e}{c} b \frac{h^2}{m^2 v_z^2} l^2,$$

and since $\mathfrak{M} = \frac{eh}{4\pi m_0 c}$:

$$\frac{\lambda l}{d} \gg \left(\frac{m_0}{m} \right)^2 \mathfrak{M} b \frac{l^2}{m v_z^2}.$$

Now, since one has $\frac{d}{dt}(mv_y) = \mathfrak{M} b$, upon integrating over time $t = l / v_y$, one will get

$$mv_y = \mathfrak{M} b \frac{l}{v_z} \quad \text{and} \quad \frac{v_y}{v_z} = \mathfrak{M} b \frac{l}{m v_z^2}.$$

The deviation D is then:

$$D = \frac{v_y}{v_z} l = \mathfrak{M} b \frac{l^2}{m v_z^2}.$$

Finally:

$$\frac{\lambda l}{d} \gg \left(\frac{m_0}{m} \right)^2 D.$$

If one supposes (as Pauli implicitly did) that $v \ll c$, and $m \sim m_0$ then one will have:

$$\frac{\lambda l}{d} \gg D,$$

which signifies that the effect of diffraction will mask the deviation entirely, but the same thing will no longer be true when $v \sim c$, because $(m_0 / m)^2$ will then be equal to $(1 - \beta^2)$, which is very small, and the inequality that was obtained above will no longer be true.

Therefore, three of the examples that were given by Pauli, following Bohr, established only the impossibility of measuring the proper magnetic moment in the case of $v \ll c$, but they will be insufficient for the case of $v \sim c$.

Pauli treated a fourth example in which the impossibility seems to persist even for $v \sim c$, as we shall see.

5. STOPPING AN ORIENTED ELECTRON BY A MAGNETIC FIELD GRADIENT

Suppose that some electrons move with a velocity v_z along Oz in the presence of a magnetic field that is directed along the negative z and have a gradient $\partial H_z / \partial z$ along Oz . The electrons whose magnetic moment is directed along the negative sense of Oz will be stopped after a time t such that:

$$mv_z = \mathfrak{M} \cdot \frac{\partial H_z}{\partial z} \cdot t \quad (v_z = \text{initial velocity, } m = \frac{m_0}{\sqrt{1 - \beta^2}}).$$

However, if H_x is zero on the axis then there can be no reason for the relation $\text{div } \mathbf{H} = 0$ to be true outside of the axis, and at a distance Δx from Oz , one will have:

$$H_x = \frac{\partial H_x}{\partial x} \Delta x = -\frac{\partial H_z}{\partial z} \Delta x.$$

A Larmor rotation of the electrons around the direction Ox will result, with a frequency of:

$$\nu = \frac{eH_x}{4\pi m_0 c} = \frac{\mathfrak{M} H_x}{h}.$$

Pauli said that this Larmor rotation results in an inversion of the sense of the rotation v_z after a length of time $1 / \nu$. In order for the reversal of v_z to be due to the action of the magnetic field on the proper moment, and not on the Larmor rotation, it will be necessary that:

$$t \ll \frac{h}{\mathfrak{M} H_x} \quad \text{or} \quad \mathfrak{M} \frac{\partial H_z}{\partial z} t \Delta x \ll h,$$

and since $mv_z = \mathfrak{M} (\partial H_z / \partial z) t$, one can deduce that $mv_z \Delta x \ll h$.

Now $\lambda = h / m_0 v_z$. Hence, $\Delta x \ll l$, and the diffraction phenomena will prevent the experiment from succeeding. That was Pauli's argument, which assumed explicitly that the electrons were slow. We shall now see what happens when $v \sim c$.

The formula that gives v will persist because in the proper system of the electron one will have:

$$v_0 = \frac{eH_x^{(0)}}{4\pi m_0 c} = \frac{eH_x}{4\pi m_0 c \sqrt{1-\beta^2}}, \quad \text{because} \quad H_x^{(0)} = \frac{H_x}{\sqrt{1-\beta^2}}.$$

Now, from the slowing-down of clocks $v = v_0 \sqrt{1-\beta^2} = \frac{eH_x}{4\pi m_0 c}$. Q. E. D. However, the proper moment M_z in the system of the observer must be $\mathfrak{M} \sqrt{1-\beta^2}$, because:

$$\mu_{12} = \mu_{12}^* : \quad \int \mu_{12} dv = \int \mu_{12}^* dv_0 \sqrt{1-\beta^2} = \mathfrak{M} \sqrt{1-\beta^2}.$$

One must then write:

$$\frac{d}{dt}(mv_z) = \mathfrak{M} \sqrt{1-\beta^2} \frac{\partial H_z}{\partial z},$$

and since $\sqrt{1-\beta^2}$ is always less than unity, one will have:

$$mv_z < \mathfrak{M} \cdot \frac{\partial H_z}{\partial z} \cdot t,$$

and since one must have $t = 1 / v$, one will have:

$$\frac{4\pi m_0 c}{eH_x} \gg \frac{mv_z}{\mathfrak{M} \frac{\partial H_z}{\partial z}},$$

namely:

$$\Delta x \ll \frac{4\pi m_0 c}{eH_x} \cdot \frac{eh}{4\pi m_0 c} \cdot \frac{1}{mv_z} = \lambda,$$

and once more there will be impossibility. The impossibility will then persist here for $v \sim c$.

6. MEASURING THE PROPER KINETIC MOMENT (SPIN)

If exhibiting the proper magnetic moment does not seem impossible when $v \sim c$ then the same thing will not be true for the proper kinetic moment (i.e., spin), which always seems to be masked by the uncertainty in the orbital moment, in Bohr's opinion.

Imagine a screen that is pierced by a rectangular opening whose edges are Δx and Δy . A sheaf of electrons falls upon that screen along the normal to that screen (viz., the Oz axis). The transmitted wave train has the lateral dimensions Δx and Δy , which are the uncertainties in the coordinates x and y of the corpuscle after it passes through the screen. The uncertainties in the components p_x and p_y of the quantity of motion will then be:

$$|\Delta p_x| \gg \frac{h}{\Delta x}, \quad |\Delta p_y| \gg \frac{h}{\Delta y}.$$

The orbital moment of the electron around Oz has the expression:

$$M_z = x p_y - y p_x.$$

There is then an uncertainty whose value is found between 0 and:

$$\Delta M_z = \Delta x \cdot |\Delta p_y| + \Delta y |\Delta p_x|,$$

so:

$$\Delta M_z \geq h \cdot \left(\frac{\Delta x}{\Delta y} + \frac{\Delta y}{\Delta x} \right) = h \left(\frac{[\Delta x]^2 + [\Delta y]^2}{\Delta x \Delta y} \right).$$

Since $(\Delta x - \Delta y)^2 = (\Delta x)^2 + (\Delta y)^2 - 2 \Delta x \Delta y \geq 0$, one can deduce that $\Delta M_z \geq 2h$, and one will have *a fortiori* that:

$$\Delta M_z \geq \frac{h}{4\pi}.$$

One sees that the uncertainty in M_z will be greater than $h / 4\pi$ – i.e., the value of the spin – which will make it impossible to recognize what part of the total kinetic moment is due to spin.

7. CONCLUSION

It does seem that one has a true impossibility, in principle, in the exhibition of the value of the proper magnetic moment for Dirac particles that are animated with a velocity that is close to c . Nevertheless, a practical impossibility still exists for electrons. For particles that are much lighter (e.g., J. Thibaud's electrinos), things do not have to work the same way when $v \sim c$.

However, Pauli gave another proof of the impossibility of exhibiting the magnetic moment, which is a proof of a completely different kind, because it utilizes not just the uncertainty relations, but the passage to the geometrical optics approximation in Dirac's theory. We shall next study that new proof, and in order to do that, we shall first recall how the passage to the geometrical optics approximation works in relativistic wave mechanics.

CHAPTER X

PASSING TO THE GEOMETRICAL OPTICS APPROXIMATION IN RELATIVISTIC WAVE MECHANICS

1. THE GEOMETRICAL OPTICS APPROXIMATION

In the wave mechanics of one wave function, the wave equation is:

$$\left[\left(\frac{h}{2\pi i} \cdot \frac{1}{c} \cdot \frac{\partial}{\partial t} - \frac{\varepsilon}{c} V \right)^2 - \sum_{x,y,z} \left(-\frac{h}{2\pi i} \frac{\partial}{\partial x} - \frac{\varepsilon}{c} A_x \right)^2 \right] \psi = m_0^2 c^2 \psi,$$

which reduces to:

$$\Delta \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = \frac{4\pi^2}{h^2} m_0^2 c^2 \psi$$

in the absence of a field. Since ψ is a complex function, we can set:

$$\psi = a(x, y, z, t) e^{2\pi i \varphi(x, y, z, t)} = a(x, y, z, t) e^{\frac{2\pi i}{h} S(x, y, z, t)},$$

in which a is the amplitude, and $2\pi\varphi$ is the phase; on the other hand, $S = h\varphi$.

If we substitute that form into the wave equation then we will get a complex equation in which, after suppressing the common factor of $e^{\frac{2\pi i}{h} S}$, we can separate the real and imaginary terms. That will give:

$$(X.a) \quad \left(\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V \right)^2 - \sum_{x,y,z} \left(\frac{\partial S}{\partial x} + \frac{\varepsilon}{c} A_x \right)^2 - \frac{h^2}{4\pi^2} \frac{\square a}{a} = m_0^2 c^2, \quad (J)$$

$$(X.b) \quad \frac{1}{c} \frac{\partial}{\partial t} \left[a^2 \left(\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V \right) \right] + \text{div} \left[a^2 \left(\text{grad } S + \frac{\varepsilon}{c} \mathbf{A} \right) \right] = 0. \quad (S)$$

Equation (X.b) expresses rigorously the conservation of the particle number, because in the relativistic wave mechanics of one wave function, the components of the density-flux quadri-vector are given by:

$$\rho = -\frac{h}{4\pi i m_0 c^2} \left(\psi \frac{\partial \psi^*}{\partial t} - \psi^* \frac{\partial \psi}{\partial t} \right) - \frac{\varepsilon^2}{m_0 c^2} V \psi \psi^2,$$

$$\mathbf{f} = \frac{h}{4\pi i m_0 c^2} (\psi \text{grad } \psi^* - \psi^* \text{grad } \psi) - \frac{\mathcal{E}^2}{m_0 c^2} \mathbf{A} \psi \psi^2,$$

which can be written:

$$\rho = \frac{1}{m_0 c} a^2 \left(\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\mathcal{E}}{c} V \right), \quad \mathbf{f} = - \frac{a^2}{m_0} \left(\text{grad } S + \frac{\mathcal{E}}{c} \mathbf{A} \right),$$

or, upon setting:

$$\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\mathcal{E}}{c} V = \pi_4, \quad -\frac{\partial S}{\partial x} - \frac{\mathcal{E}}{c} A_x = \pi_x, \dots,$$

one can write:

$$\rho = \frac{1}{m_0 c} a^2 \pi_4, \quad \mathbf{f} = \frac{a^2}{m_0} \boldsymbol{\pi},$$

and equation (X.b), which one can write as:

$$\frac{1}{c} \frac{\partial}{\partial t} (a^2 \pi_4) + \text{div} (a^2 \boldsymbol{\pi}) = 0,$$

will be equivalent to $\frac{\partial \rho}{\partial t} + \text{div } \mathbf{f} = 0$. It will express the conservation of the probability of presence (or of particle number in the statistical case).

Return to equation (X.a), which one can write as:

$$(X.c) \quad \pi_4^2 - \boldsymbol{\pi}^2 - \frac{h^2}{4\pi^2} \frac{\square a}{a} = m_0^2 c^2.$$

It links the variations of the amplitude a to those of the phase S . One says that the geometrical optics approximation is valid when one can neglect the terms in h^2 in that equation. One proves that this condition is realized when the conditions of propagation are such that the amplitude will vary only slightly at the scale of the wave length. One will then have:

$$\pi_4^2 - \boldsymbol{\pi}^2 = m_0^2 c^2,$$

and that relation is the one that exists in relativistic point mechanics, because in that mechanics (i.e., Einsteinian dynamics), one will have:

$$\frac{W}{c} = \frac{\partial S}{\partial t} = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + \mathcal{E}V, \quad \mathbf{p} = - \text{grad } S = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}} + \frac{\mathcal{E}}{c} \mathbf{A},$$

so:

$$(X.d) \quad \pi_4 = \frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V = \frac{m_0 c^2}{\sqrt{1-\beta^2}}, \quad \boldsymbol{\pi} = -\text{grad } S - \frac{\varepsilon}{c} \mathbf{A} = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}},$$

which indeed implies that:

$$\pi_4^2 - \boldsymbol{\pi}^2 = m_0^2 c^2.$$

In that case, the phases of the wave are represented by a complete integral of the Jacobi equation:

$$\left(\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V \right)^2 - \sum_{x,y,z} \left(\frac{\partial S}{\partial x} + \frac{\varepsilon}{c} \mathbf{A} \right)^2 = m_0^2 c^2,$$

and the motion is defined at each point of space-time by formulas (X.d), which give π_4 and $\boldsymbol{\pi}$ as functions of $S(x, y, z, t)$, $V(x, y, z, t)$, and $\mathbf{A}(x, y, z, t)$, which are assumed to be continuous.

From the viewpoint of wave mechanics, one can represent what happens in the following fashion: The conditions of propagation vary only slightly at the scale of the wave length, so one can consider regions that are very large with respect to the wave length, along with potentials V and \mathbf{A} that vary only slightly inside of them. In such a region, there can exist a group of waves whose dimensions are very large compared to the wave length, and which are meanwhile almost monochromatic with frequency $\nu = \frac{\partial S}{\partial t} \frac{1}{h}$

and wave length $\lambda = \frac{h}{[\text{grad } S]}$. The quantities π_4 and $\boldsymbol{\pi}$ are defined for that wave group by the formulas:

$$\pi_4 = \frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V, \quad \boldsymbol{\pi} = -\text{grad } S - \frac{\varepsilon}{c} \mathbf{A},$$

with values of V and \mathbf{A} that are almost constant *locally*. The equation $\frac{\partial \rho}{\partial t} + \text{div } \mathbf{f} = 0$ then shows that the wave group displaces with a velocity $[\mathbf{f}] / \rho$ that is equal to:

$$\frac{\boldsymbol{\pi}}{\pi_4} c = \frac{\frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}}}{\frac{m_0 c}{\sqrt{1-\beta^2}}} c = \mathbf{v}.$$

One is then dealing with a group of waves – or “globule” – of probability that slides along one of the trajectories that are predicted by classical point mechanics with a velocity of \mathbf{v} , and has performed the transition from wave mechanics to point mechanics, in the approximation of geometrical optics.

In summary, geometrical optics, and in turn, point mechanics are valid as long as one can obtain a Jacobi equation by neglecting the terms of order h^2 , so the equation of

continuity then proves that one can consider wave groups to displace along one of the trajectories of point mechanics with a corresponding velocity.

2. THE WKB METHOD OF SUCCESSIVE APPROXIMATIONS

The Wentzel-Kramers-Brillouin method proceeds by successive approximation with respect to the very small quantity $h / 2\pi i$. In order to do that, one starts with one of the following developments:

$$\psi = e^{\frac{2\pi i}{h} \left[S_0 + \frac{h}{2\pi i} S_1 + \dots + \left(\frac{h}{2\pi i} \right)^n S_n + \dots \right]},$$

or

$$\psi = \left[a_0 + \frac{h}{2\pi i} a_1 + \dots + \left(\frac{h}{2\pi i} \right)^n a_n + \dots \right] e^{\frac{2\pi i}{h} S_0},$$

between which exist the relations:

$$a_0 = e^{S_1}, \quad a_1 = S_2 e^{S_1}, \quad a_2 = \left(\frac{S_2}{2} + S_3 \right) e^{S_1}, \quad \dots$$

The zero-order approximation (in $h / 2\pi i$) consists of keeping only the term in a_0 (or S_0). The first-order approximation consists of keeping a_0 and a_1 (or S_0 and S_1), ...

In the preceding method, we write $\psi = a e^{\frac{2\pi i}{h} S}$, in which a and S are the modulus and argument of ψ , resp. Here (this remark will be important in what follows), the a_0, a_1 are less well-defined. For example, if one has $a_i = \alpha_i e^{i\beta_i}$ (in which α_i and β_i are the modulus and argument, resp., of a_i) then one can introduce β_0 into the exponential factor by setting $S = S_0 + \frac{h}{2\pi i} \beta_0$, and set:

$$\psi = \left[\alpha_0 + \frac{h}{2\pi i} \alpha_1 e^{i(\beta_1 - \beta_0)} + \dots + \left(\frac{h}{2\pi i} \right)^n \alpha_n e^{i(\beta_n - \beta_0)} + \dots \right] e^{\frac{2\pi i}{h} S},$$

so there will be a certain indeterminacy in the form of the development, above all, if β_0 is large enough that $\frac{h}{2\pi i} \beta_0$ has an order of magnitude that is between 1 and $\frac{h}{2\pi i}$. In other words, if $S_1 = S'_1 + i S''_1$ has an imaginary part S''_1 that is large enough that $\frac{h}{2\pi i} \cdot S_1$ will have an order of magnitude that is between 1 and $\frac{h}{2\pi i}$, and one does not even know if one can write:

$$a_0 = e^{S'_1}$$

or

$$a_0 = e^{S'_1 + iS''_1}$$

or even:

$$a_0 = e^{S'_1 + i\eta S''_1},$$

with

$$0 < \eta < 1.$$

Be that as it may, and upon assuming that those indeterminacies are not troublesome, we shall substitute the development in a_i in the equation of propagation, and separate the corresponding terms in various powers of $h / 2\pi i$.

The *zero-order* equation in $h / 2\pi i$ is:

$$\left(\frac{1}{c} \frac{\partial S_0}{\partial t} - \frac{\mathcal{E}}{c} V \right)^2 - \sum_{i=1}^3 \left(\frac{\partial S_0}{\partial x_i} + \frac{\mathcal{E}}{c} A_i \right)^2 = m_0^2 c^2.$$

That the Jacobi equation: It shows that $2\pi S_0 / h$ is the phase, in the geometrical optics approximation.

The equation of order *one* in $h / 2\pi i$ gives:

$$\frac{1}{c} \frac{\partial}{\partial t} [\pi_4^{(0)} a_0^2] + \text{div} [\boldsymbol{\pi}^{(0)} a_0^2] = 0,$$

with:

$$\pi_4^{(0)} = \frac{1}{c} \frac{\partial S_0}{\partial t} - \frac{\mathcal{E}}{c} V, \quad \boldsymbol{\pi} = -\text{grad } S_0 - \frac{\mathcal{E}}{c} \mathbf{A},$$

and that equation always expresses the conservation of the probability whose density is a_0^2 , and whose velocity is $\frac{\boldsymbol{\pi}^{(0)} c}{\pi_4^{(0)}}$ in the geometrical optics approximation.

The following equations will permit one to calculate a_1, a_2, \dots as functions of S_0 and a_0 . Those quantities can be neglected in the geometrical optics approximation. If they are not negligible then diffraction phenomena will appear.

The two methods that we just employed in order to pass to the geometrical optics approximation in wave mechanics seem simple in non-relativistic wave mechanics, and as we have just seen, in relativistic wave mechanics with one ψ . Things are completely different in Dirac's theory, in which the question much more complicated. We shall now study it by successively imagining the two methods that we just presented.

3. PASSING TO GEOMETRICAL OPTICS IN DIRAC'S THEORY

In order to effect the passage to geometrical optics, we first employ the same method as in relativistic wave mechanics.

We start from the second-order equations of Dirac theory, which are:

$$\left[P_4^2 - \sum_{i=1}^3 P_i^2 - m_0^2 c^2 + \frac{\varepsilon}{c} \frac{h}{2\pi i} (\alpha_1 h_x + \alpha_2 h_y + \alpha_3 h_z) - \frac{\varepsilon}{c} \frac{h}{2\pi i} (\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z) \right] \psi_k = 0, \\ (k = 1, 2, 3, 4),$$

and in order to pass to the geometrical optics approximation, we set:

$$\psi_k = a_k e^{\frac{2\pi i S}{h}} \quad (k = 1, 2, 3, 4),$$

in which S is a function of phase that is common to the four ψ_k and rapidly-varying at the scale of wave length, while the a_k are amplitudes that must be slowly-varying variables in the geometrical optic approximation. We shall substitute that form for the ψ_k in the second-order symbolic equation, and upon separating the real and imaginary terms, we shall seek to deduce, on the one hand, a Jacobi (J), and on the other, a continuity equation (C). However, in order to be able to separate the real and pure imaginary terms with any confidence, one must take some precautions. In order to obtain that separation after the introduction of the form that was assumed for the ψ_k into the second-order symbolic equation, we will apply the operator α_4 to that equation, and then we will multiply by a_k^* and sum over k . We can then separate the real and imaginary terms, and upon setting:

$$\pi_4 = \frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V, \quad \boldsymbol{\pi} = -\text{grad } S - \frac{\varepsilon}{c} \mathbf{A},$$

we will get, for the real terms:

$$\begin{aligned} & (\pi_4^2 - \pi^2 - m_0^2 c^2) \sum_{k=1}^4 a_k^* \alpha_4 a_k + \frac{h}{2\pi i} \sum_{k=1}^4 \left[\left(a_k^* \alpha_4 \frac{1}{c} \frac{\partial a_k}{\partial t} - \frac{1}{c} \frac{\partial a_k^*}{\partial t} \alpha_4 a_k \right) \right. \\ & \left. + \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial x} - \frac{\partial a_k}{\partial x} \alpha_4 a_k \right) \pi_x + \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial y} - \frac{\partial a_k}{\partial y} \alpha_4 a_k \right) \pi_y + \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial z} - \frac{\partial a_k}{\partial z} \alpha_4 a_k \right) \pi_z \right] \\ & - \frac{h^2}{2\pi i} \frac{1}{2} \sum_{k=1}^4 (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k) \\ & = - \frac{\varepsilon}{c} \frac{h}{2\pi i} \sum_{k=1}^4 a_k^* (h_x i \alpha_1 \alpha_4 + h_y i \alpha_2 \alpha_4 + h_z i \alpha_3 \alpha_4) a_k \\ & - \frac{\varepsilon}{c} \frac{h}{2\pi i} \sum_{k=1}^4 a_k^* (H_x i \alpha_2 \alpha_3 \alpha_4 + H_y i \alpha_3 \alpha_1 \alpha_4 + H_z i \alpha_1 \alpha_2 \alpha_4) a_k. \end{aligned}$$

That gives the Jacobi equation (J).

As for the imaginary terms, they will give the equation (C):

$$\frac{1}{c} \frac{\partial}{\partial t} \left(\pi_4 \sum_{k=1}^4 a_k^* \alpha_4 a_k \right) + \operatorname{div} \left(\boldsymbol{\pi} \sum_{k=1}^4 a_k^* \alpha_4 a_k \right) + \frac{h}{2\pi i} \sum_{k=1}^4 (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k) = 0.$$

We shall study the first equation first. In the geometrical optics approximation, the wave ψ will reduce to a wave packet that is almost monochromatic in a region where the field is reasonably constant. It can be assimilated into a “local” monochromatic plane wave that corresponds to a certain velocity βc , and one will have:

$$\sum_{k=1}^4 a_k^* \alpha_4 a_k = \rho \sqrt{1 - \beta^2},$$

in which ρ is the probability density, and as a result:

$$\int \sum_{k=1}^4 a_k^* \alpha_4 a_k d\tau = \sqrt{1 - \beta^2}, \quad \text{because} \quad \int \rho d\tau = 1.$$

Upon integrating the first equation over D , we will get the equation:

$$\pi_4^2 - \boldsymbol{\pi}^2 + F(a_k) + G(a_k) = m_0^2 c^2 - \frac{2}{m_0 c \sqrt{1 - \beta^2}} \{(\mathbf{H} \cdot \boldsymbol{\mathfrak{M}}) + (\mathbf{h} \cdot \boldsymbol{\mathfrak{P}})\},$$

in which $\boldsymbol{\mathfrak{M}}$ and $\boldsymbol{\mathfrak{P}}$ are the proper magnetic moment and the proper electric moment of the particle, respectively. By definition, we will have:

$$\begin{aligned} G(a_k) &= -\frac{h^2}{8\pi^2} \frac{\sum_{k=1}^4 (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k)}{\sqrt{1 - \beta^2}} \\ F(a_k) &= \frac{h}{2\pi i} \frac{\sum_{k=1}^4 \left[\left(a_k^* \alpha_4 \frac{1}{c} \frac{\partial a_k}{\partial t} - \frac{1}{c} \frac{\partial a_k^*}{\partial t} \alpha_4 a_k \right) \pi_4 + \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial x} - \frac{\partial a_k^*}{\partial x} \alpha_4 a_k \right) \pi_x \right.}{\sqrt{1 - \beta^2}} \\ &\quad \left. + \frac{\left(a_k^* \alpha_4 \frac{\partial a_k}{\partial y} - \frac{\partial a_k^*}{\partial y} \alpha_4 a_k \right) \pi_y + \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial z} - \frac{\partial a_k^*}{\partial z} \alpha_4 a_k \right) \pi_z}{\sqrt{1 - \beta^2}} \right] \\ &= \frac{h}{2\pi i} \frac{\sum_{k=1}^4 \sum_{j=1}^4 \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial x_j} - \frac{\partial a_k^*}{\partial x_j} \alpha_4 a_k \right) \pi_j}{\sqrt{1 - \beta^2}}. \end{aligned}$$

We shall suppose that $F(a_k)$ and $G(a_k)$ are negligible in the geometrical optics approximation. That seems quite reasonable for $G(a_k)$, which has order h^2 ; it is less obvious for $F(a_k)$, which has order h . We shall return much later to these hypotheses in order to justify them; for the moment, we shall just assume them. We will then get the Jacobi equation:

$$\pi_4^2 - \pi^2 = m_0^2 c^2 + 2\eta U m_0,$$

with the definitions:

$$\eta = \frac{1}{\sqrt{1-\beta^2}}, \quad U = -(\mathfrak{M} \cdot \mathbf{H}) - (\mathfrak{P} \cdot \mathbf{h}).$$

U is the potential energy that is due to the action of the field on the proper moments.

Upon writing the right-hand side of the Jacobi equation in the form:

$$m_0'^2 c^2 = m_0^2 c^2 + 2m_0 \eta U,$$

we will define a variable proper mass that is precisely analogous to the one that Weyssenhoff introduced into his dynamics of the spinning particle. We should have expected that coincidence. The mass m_0 is slowly-varying at the scale of wave length.

In the preceding formula, one can remark that ηU is a relativistic invariant. That might be surprising, since one would be tempted to give U that relativistic variance of energy, but ηU would not be an invariant then. However, in reality, U does not have the relativistic variance of energy, although it does the same dimension. Indeed, if F_{ik} and m_{ik} denote the components of the *electromagnetic field* tensor and those of the *densities* of proper magnetic moment and electric moment, resp., then:

$$U = -\frac{1}{2} \int \sum_{i,k} F_{ik} m^{ik} d\tau = -\frac{1}{2} \int \sum_{i,k} F_{ik} m^{ik} d\tau_0 \sqrt{1-\beta^2} = U_0 \sqrt{1-\beta^2} = \frac{U_0}{\eta},$$

in which U_0 is the value of U in the proper system. Hence, the product ηU is indeed and invariant.

It seems very probable that in order to obtain a point mechanics for spinning particles, one must suppose that $2m_0 U_0$ is very small compared to $m_0^2 c^2$; i.e., that the proper energy U_0 that is due to the action of a field on the proper moments is very small compared to the energy of the mass $m_0 c^2$. We will then make the hypothesis:

$$\frac{U_0}{m_0 c^2} = \frac{\eta U}{m_0 c^2} = K$$

is very small compared to unity without also supposing that K is absolutely negligible. We can then set, approximately:

$$m_0'^2 c^2 \sim \left(m_0 c + \frac{U_0}{c} \right)^2 = \left(m_0 c + \frac{\eta U}{c} \right)^2,$$

so:

$$m_0'^2 = m_0 + \frac{U_0}{c^2} = m_0 + \frac{\eta U}{c^2}.$$

It will then be natural to seek to develop the point mechanics of the spinning electron by setting the Lagrange function equal to:

$$\mathcal{L} = -m_0'^2 c^2 \sqrt{1-\beta^2} - \varepsilon V + \frac{\varepsilon}{c} (\mathbf{v} \cdot \mathbf{A});$$

i.e., upon replacing m_0 with m_0' in the usual expression for the Lagrange function in the relativistic dynamics of an electron. One can then write:

$$\mathcal{L} = -m_0'^2 c^2 \sqrt{1-\beta^2} - U - \varepsilon V + \frac{\varepsilon}{c} (\mathbf{v} \cdot \mathbf{A}),$$

which is an expression that is indeed such that $\int \mathcal{L} dt$ is a relativistic invariant, since:

$$\int U dt = \int U_0 dt_0 = \int \left(\frac{U_0 dt}{\sqrt{1-\beta^2}} - \sum_{i=1}^3 \frac{U_0}{c^2 \sqrt{1-\beta^2}} v_i dx_i \right)$$

is obviously invariant, and the integrals are taken along the world-lines of the particle.

The classical formulas of analytical dynamics $p_i = \partial \mathcal{L} / \partial \dot{q}_i$ give us:

$$\mathbf{p} = -\text{grad } S = \frac{m_0' \mathbf{v}}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} \mathbf{A},$$

so:

$$\boldsymbol{\pi} \equiv \mathbf{p} - \frac{\varepsilon}{c} \mathbf{A} = \frac{m_0' \mathbf{v}}{\sqrt{1-\beta^2}} = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}} + \frac{U_0 \mathbf{v}}{c^2 \sqrt{1-\beta^2}} = \boldsymbol{\pi}^{(0)} + \frac{U_0 \mathbf{v}}{c^2 \sqrt{1-\beta^2}}.$$

As for the energy, it is given by formula:

$$W = \frac{\partial S}{\partial t} = \sum \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \mathcal{L} = \frac{m_0' c^2}{\sqrt{1-\beta^2}} + \varepsilon V,$$

from which, one will deduce that:

$$\pi_4 = -\frac{1}{c} \frac{\partial S}{\partial t} - \frac{\varepsilon}{c} V = \frac{m_0' c^2}{\sqrt{1-\beta^2}} = \frac{m_0 c}{\sqrt{1-\beta^2}} + \frac{U_0}{c \sqrt{1-\beta^2}} = \pi_4^{(0)} \frac{U_0}{c \sqrt{1-\beta^2}}.$$

$c \pi_4^{(0)} = \frac{m_0 c^2}{\sqrt{1-\beta^2}}$ and $\boldsymbol{\pi}^{(0)} = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}}$ denote the energy and quantity of motion,

respectively, of a particle of proper mass m_0 and velocity \mathbf{v} in the absence of that field.

One immediately verifies the Jacobi equation (J), because:

$$\pi_4^2 - \boldsymbol{\pi}^2 = m_0'^2 c^2 \sim m_0^2 c^2 + 2m_0 U_0 = m_0^2 c^2 + 2m_0 \eta U,$$

with the hypothesis that $k \ll 1$.

If we now write the Lagrange equations:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i}$$

then with the form that we assumed for \mathcal{L} , we will get:

$$\frac{\partial \mathbf{p}}{\partial t} = \mathbf{f} - \text{grad } U,$$

in which \mathbf{f} is the Lorentz force, and $-\text{grad } U$ is the force that is exerted by the gradient of the electromagnetic field in the proper moment.

Upon comparing the value of the force $-\text{grad } U$ with the uncertainty in the Laplacian [*sic*] force that is contained in \mathbf{f} in the case that was envisioned at the beginning of Chapter IX, one will easily recover the formula $\Delta x < h / m_0 v$, and the consequences of it that we inferred (viz., the impossibility of exhibiting the proper magnetic moment unless $v \sim c$).

Now return to equation (C):

$$0 = \frac{1}{c} \frac{\partial}{\partial t} \left(\pi_4 \sum_{k=1}^4 a_k^* \alpha_4 a_k \right) + \text{div} \left(\boldsymbol{\pi} \sum_{k=1}^4 a_k^* \alpha_4 a_k \right) + \frac{h}{2\pi i} (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k).$$

In order to see that this represents the equation of continuity, we first make a less rigorous argument: We neglect the last term and let π_4 and $\boldsymbol{\pi}$ coincide with $\pi_4^{(0)}$ and $\boldsymbol{\pi}^{(0)}$, resp., which will give us:

$$\frac{1}{c} \frac{\partial}{\partial t} \left(\frac{m_0 c}{\sqrt{1-\beta^2}} \cdot \rho \sqrt{1-\beta^2} \right) + \text{div} \left(\frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}} \rho \sqrt{1-\beta^2} \right) = 0,$$

or after dividing by the constant m_0 :

$$\frac{\partial \rho}{\partial t} + \text{div} (\rho \mathbf{v}) = 0,$$

which is indeed the equation of continuity.

However, in reality, $\pi_4 = \frac{m'_0 c}{\sqrt{1-\beta^2}}$, $\boldsymbol{\pi} = \frac{m'_0 \mathbf{v}}{\sqrt{1-\beta^2}}$, and one must take into account the

last term, so:

$$\frac{\partial \rho}{\partial t} (\rho m'_0) + \operatorname{div}(\rho \mathbf{v} m'_0) + \frac{h}{2\pi i} \sum_{k=1}^4 (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k) = 0,$$

One can write:

$$\frac{\partial \rho}{\partial t} (\rho m'_0) + \operatorname{div}(\rho \mathbf{v} m'_0) = m'_0 \left[\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) \right] + \rho \frac{dm'_0}{dt},$$

in which $\frac{dm'_0}{dt} = \frac{\partial m'_0}{\partial t} + \mathbf{v} \operatorname{grad} m'_0$ is the total derivative of m'_0 along the trajectory. The equation of continuity will then be verified if:

$$\frac{1}{c^2} \frac{dU_0}{dt} = \frac{dm'_0}{dt} = - \frac{h}{2\pi i} \sum_{k=1}^4 (a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k).$$

The last term in equation (C), which can initially appear to be a parasitic term, actually serves to compensate for the variations of m'_0 along the trajectory. Later on, we shall recover a proof of the equation of continuity.

4. THE WKB METHOD IN DIRAC'S THEORY

We have carried out the passage to geometrical optics by starting from the second-order equation. One can also start from the first-order equations and apply the WKB method. That is what Pauli did in a very interesting paper [Helvetica Physica Acta (1932), pp. 179], although we shall not adopt all of its conclusions. We shall, however, reproduce its argument.

We start with the Dirac's first-order equations:

$$\left(\frac{1}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} - \frac{\boldsymbol{\varepsilon}}{c} V \right) \psi_k + \sum_{i=1}^3 \left(-\alpha_i \frac{h}{2\pi i} \frac{\partial}{\partial x_i} - \frac{\boldsymbol{\varepsilon}}{c} A_i \right) \psi_k = m_0 c \alpha_4 \psi_k \quad (k = 1, 2, 3, 4).$$

With a slight change of notations, with Pauli, we write the development of the ψ_k as:

$$\psi_k = b_k e^{\frac{2\pi i}{h} S_0} = \left[b_k^{(0)} + \frac{h}{2\pi i} b_k^{(1)} + \dots + \left(\frac{h}{2\pi i} \right)^n b_k^{(n)} + \dots \right] e^{\frac{2\pi i}{h} S_0},$$

in which the b_k on right-hand side are developed in powers of the very small quantity $h / 2\pi i$ (conforming to the principles of the WKB method), and in which S_0 denotes the

Jacobi function for a particle *without spin* that is subject to the electromagnetic field under consideration; i.e., the function:

$$S_0 = \int \left[\left(\pi_4^{(0)} + \frac{\mathcal{E}}{c} V \right) c dt - \sum_{i=1}^3 \left(\pi_i^{(0)} + \frac{\mathcal{E}}{c} A_i \right) dx_i \right],$$

in which the integral is taken along the ray-trajectory. One must indeed note that in the preceding formula, it is $\pi_4^{(0)}$ and $\pi_i^{(0)}$, and not π_4 and $\boldsymbol{\pi}$ that appear in the right-hand side. One must also note that the quantities:

$$\pi_4^{(0)} = \frac{1}{c} \frac{\partial S_0}{\partial t} - \frac{\mathcal{E}}{c} V(x, y, z, t), \quad \pi_i^{(0)} = -\frac{\partial S_0}{\partial x_i} - \frac{\mathcal{E}}{c} A_i(x, y, z, t)$$

are well-defined functions of x, y, z, t that are slowly-varying at the scale of the electromagnetic field when the geometrical optics approximation is valid. Hence, in a permanent field, according to the Pauli viewpoint, one will have:

$$c \pi_4^{(0)} = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} = W - eV(x, y, z, t),$$

in which W is the constant value of the energy.

Upon substituting the form that was assumed for the ψ_k in the Dirac equations, and calculating the successive approximations in the powers of $h / 2\pi i$, one will get the equations:

$$(X.e) \quad \begin{cases} (a) \quad \left[\pi_4^{(0)} + \sum_{i=1}^3 \alpha_i \pi_i^{(0)} - \alpha_4 m_0 c \right] b_k^{(0)} = 0, \\ (b) \quad \left[\pi^{(0)} + \sum_{i=1}^3 \alpha_i \pi_i^{(0)} - \alpha_4 m_0 c \right] b_k^{(1)} = - \left(\frac{1}{c} \frac{\partial b_k^{(0)}}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial b_k^{(0)}}{\partial t} \right). \end{cases}$$

Since the first equation is homogeneous, it will admit non-zero solutions only if the determinant is zero, and that condition is equivalent to the relation:

$$(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 = m_0^2 c^2,$$

which is indeed verified. It will then result from the well-known theory of monochromatic plane waves in Dirac's theory that the $b_k^{(0)}$ are the amplitudes of the ψ_k in the uniform, rectilinear motion that corresponds to the energy $c \pi_4^{(0)}$ and the quantity of motion $\boldsymbol{\pi}^{(0)}$. However, as is well-known, each state of uniform, rectilinear motion will correspond to two possible spin states. If we let A_k and B_k denote the amplitudes (i.e., functions of $\pi_4^{(0)}$ and $\boldsymbol{\pi}^{(0)}$, and in turn, slowly-varying variables at the large scale) that relate to the two spin states, and whose expressions we know, then we will have:

$$b_k^{(0)} = C_1(x, y, z, t) A_k(\boldsymbol{\pi}_4^{(0)}, \boldsymbol{\pi}^{(0)}) + C_2(x, y, z, t) B_k(\boldsymbol{\pi}_4^{(0)}, \boldsymbol{\pi}^{(0)}),$$

with:

$$A_1 = -\frac{\pi_z^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad A_2 = -\frac{\pi_x^{(0)} - i\pi_y^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad A_3 = 1, \quad A_4 = 0,$$

$$B_1 = -\frac{\pi_x^{(0)} + i\pi_y^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad B_2 = \frac{\pi_z^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad B_3 = 0, \quad B_4 = 1.$$

C_1 and C_2 are slowly-varying variables at the large scale.

One then sees that the equations in $b_k^{(1)}$, which are linear and have a zero determinant, can admit non-zero solutions for $b_k^{(1)}$ only if the following conditions are realized:

$$\sum_{k=1}^4 A_k^* \left(\frac{1}{c} \frac{\partial b_k^{(0)}}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial b_k^{(0)}}{\partial x_i} \right) = 0, \quad \sum_{k=1}^4 B_k^* \left(\frac{1}{c} \frac{\partial b_k^{(0)}}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial b_k^{(0)}}{\partial x_i} \right) = 0.$$

Those are the *Pauli conditions*. If one is given the expressions above for A_k and B_k then it will be easy to write them out explicitly; for example, in the case where the $\pi_4^{(0)}$ remain constant.

$$(X.f)(a) \quad 2 \left[\pi_4^{(0)} \frac{1}{c} \frac{\partial C_1}{\partial t} + \pi_x^{(0)} \frac{\partial C_1}{\partial x} + \pi_y^{(0)} \frac{\partial C_1}{\partial y} + \pi_z^{(0)} \frac{\partial C_1}{\partial z} \right]$$

$$+ \pi_z^{(0)} \left(\frac{\partial C_2}{\partial x} - \frac{\partial C_2}{\partial y} \right) - (\pi_x^{(0)} - i\pi_y^{(0)}) \frac{\partial C_2}{\partial z}$$

$$= -C_1 \operatorname{div} \boldsymbol{\pi}^{(0)} + i C_1 \left(\frac{\partial \pi_y^{(0)}}{\partial x} - \frac{\partial \pi_x^{(0)}}{\partial y} \right) - i C_2 \left[\left(\frac{\partial \pi_x^{(0)}}{\partial y} - \frac{\partial \pi_y^{(0)}}{\partial x} \right) + i \left(\frac{\partial \pi_x^{(0)}}{\partial z} - \frac{\partial \pi_z^{(0)}}{\partial x} \right) \right]$$

and

$$(X.f)(b) \quad 2 \left[\pi_4^{(0)} \frac{1}{c} \frac{\partial C_2}{\partial t} + \pi_x^{(0)} \frac{\partial C_2}{\partial x} + \pi_y^{(0)} \frac{\partial C_2}{\partial y} + \pi_z^{(0)} \frac{\partial C_2}{\partial z} \right]$$

$$+ \pi_z^{(0)} \left(\frac{\partial C_1}{\partial x} + \frac{\partial C_1}{\partial y} \right) - (\pi_x^{(0)} + i\pi_y^{(0)}) \frac{\partial C_2}{\partial z}$$

$$= -C_2 \operatorname{div} \boldsymbol{\pi}^{(0)} - i C_2 \left(\frac{\partial \pi_y^{(0)}}{\partial x} - \frac{\partial \pi_x^{(0)}}{\partial y} \right) - i C_1 \left[\left(\frac{\partial \pi_x^{(0)}}{\partial y} - \frac{\partial \pi_y^{(0)}}{\partial x} \right) + i \left(\frac{\partial \pi_x^{(0)}}{\partial z} - \frac{\partial \pi_z^{(0)}}{\partial x} \right) \right].$$

It is important to remark that the first bracket in these two equations can be written $m_0 c \eta dC_i / dt$ (with $i = 1, 2$), where the derivative d / dt is taken along the motion of the particle that is defined by the vector $\boldsymbol{\pi}^{(0)}$.

The equations for the Pauli condition are two first-order partial differential equations that determine the functions $C_1(x, y, z, t)$ and $C_2(x, y, z, t)$ when one knows the initial velocities. The variation at the large scale of the A_k and B_k , by the intermediary of the

$\pi_4^{(0)}$ and $\boldsymbol{\pi}^{(0)}$, and the variations of the functions C_1 and C_2 , express the way that the portion of the monochromatic plane wave that represents the motion of a wave packet at the small scale will be transformed under the influence of the electromagnetic field when the wave packet progresses.

Multiply the equations for the Pauli conditions by C_1^* and C_2^* , respectively, and the conjugate equations by C_1 and C_2 , respectively, and add them. If we take into account the Hermiticity of the α_i then we will get:

$$\frac{1}{c} \frac{\partial}{\partial t} \sum_{k=1}^4 b_k^{(0)*} b_k^{(0)} - \sum_{i=1}^3 \frac{\partial}{\partial x_i} \sum_{k=1}^4 b_k^{(0)*} \alpha_i b_k^{(0)} = 0,$$

and since one has:

$$\rho = \sum_{k=1}^4 b_k^{(0)*} b_k^{(0)}, \quad \mathbf{f} = \rho \mathbf{u} = \sum_{k=1}^4 b_k^{(0)*} (-c \alpha_i) b_k^{(0)}$$

(at least in the zero-order approximation), one will get the equation of continuity:

$$\frac{\partial \rho}{\partial t} + \text{div} (\rho \mathbf{u}) = 0.$$

Of course, one can study the determination of the $b_k^{(2)}$, $b_k^{(3)}$, ... in an analogous manner.

Just as the A_k and B_k do not in any way depend upon the action of the electromagnetic field in the proper moments of the electron, the same thing will be true for the $b_k^{(0)}$. (This is a point that we need to examine more closely.) Hence, the zero-order approximation will not at all involve the action of the field on the proper moment of the electron, and the equation of continuity, with the given expressions for ρ and $\rho \mathbf{u}$, shows that the motion is the same as if the proper moments did not exist in that approximation. Pauli concluded that when point mechanics is valid for a Dirac electron that will be the same thing as the electron having no spin. In order to find appreciable spin effects, one will be obliged to take into account the $b_k^{(1)}$. However, it is well-known that if one is obliged to take into account the $b_k^{(1)}$ then diffraction effects will appear, and point mechanics will cease to be valid. Pauli then thought that it would be impossible to exhibit the existence of the proper magnetic moment of an electron by experiments in which the concepts of point mechanics are valid; i.e., in which one can attribute well-defined trajectories to the electrons.

The conclusion to which we, with Pauli, then arrive, according to which point mechanics of a Dirac electron is identical to point mechanics of an electron without spin, is obviously in complete contradiction to the ideas that we have developed above, as well as with Weysenhoff's theory. We must then subject the question of a deeper examination.

5. APPLYING PAULI'S LINE OF REASONING WHEN ONE STARTS WITH WEYSENHOFF'S POINT MECHANICS

In the course of his arguments, Pauli started with the point mechanics of an electron without spin in his calculation of the approximations. From our viewpoint, it is natural to repeat the same line of reasoning by starting from Weysenhoff's point mechanics, which we recovered above. From the outset, we shall then envision an ensemble of trajectories of the same class (i.e., ones that correspond to the same Jacobi function) in the context of Weysenhoff's point mechanics. Of course, those trajectories are distinct from the ones that were envisioned in Pauli's argument, since they are subject to the influence of the action of electromagnetic fields upon the proper moments. The wave ψ that is associated with the class of trajectories considered can be written in the form:

$$\psi_k = a_k e^{\frac{2\pi i S}{h}} = \left[a_k^{(0)} + \frac{h}{2\pi i} a_k^{(1)} + \left(\frac{h}{2\pi i} \right)^2 a_k^{(2)} + \dots \right] e^{\frac{2\pi i S}{h}},$$

in which S is the Jacobi function that is defined by:

$$S = \int \left[\left(\frac{m_0 c^2}{\sqrt{1-\beta^2}} + \varepsilon V + \frac{U_0}{\sqrt{1-\beta^2}} \right) dt - \sum_{i=1}^3 \left(\frac{m_0 v_i}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_i + \frac{U_0}{c} \cdot \frac{v_i}{\sqrt{1-\beta^2}} \right) dx_i \right],$$

in which the integral is taken along the motion.

Set:

$$S'_0 = \int \left[\left(\frac{m_0 c^2}{\sqrt{1-\beta^2}} + \varepsilon V \right) dt - \sum_{i=1}^3 \left(\frac{m_0 v_i}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_i \right) dx_i \right].$$

S'_0 does not coincide with Pauli's S_0 , since the integral is taken along a motion that is different from the one that Pauli imagined. If we further set:

$$(X.g) \quad b_k^{(j)} = a_k^{(j)} e^{\frac{2\pi i}{h} \int U dt} \quad (j = 0, 1, 2, \dots)$$

then we can write the development of the ψ_k in a form that is analogous to the one that Pauli employed:

$$\psi_k = \left[b_k^{(0)} + \frac{h}{2\pi i} b_k^{(1)} + \left(\frac{h}{2\pi i} \right)^2 b_k^{(2)} + \dots \right] e^{\frac{2\pi i S'_0}{h}}.$$

Substitute that form in the Dirac equations:

$$\left(\frac{h}{2\pi i} \frac{\partial}{\partial t} - \varepsilon V \right) \psi_k + \sum_{i=1}^3 \left(-\alpha_i \frac{h}{2\pi i} \frac{\partial}{\partial x_i} - \varepsilon A_i \right) \psi_k = \alpha_4 m_0 c^2 \psi_k.$$

Set:

$$\pi_4^{(0)} = \frac{1}{c} \frac{\partial S'_0}{\partial t} - \varepsilon V = \frac{m_0 c}{\sqrt{1-\beta^2}}, \quad \pi_i^{(0)} = -\frac{\partial S'_0}{\partial x_i} - \frac{\varepsilon}{c} A_i = \frac{m_0 v_i}{\sqrt{1-\beta^2}} \quad (i = 1, 2, 3).$$

The $\pi_4^{(0)}$ and $\boldsymbol{\pi}^{(0)}$ thus-defined are not the same as in Pauli's argument, since they correspond to Weysenhoff's point mechanics, and not to that of the electron without spin, but one will always have:

$$(\pi_4^0)^2 - (\boldsymbol{\pi}^{(0)})^2 = m_0^2 c^2.$$

The result of the substitution of the ψ_k in the Dirac equations always permits us to write the sequence of equations for the successive approximations:

$$(X.h) \quad \begin{cases} (a) \quad \left[\pi_4^{(0)} + \sum_{i=1}^4 \alpha_i \pi_i^{(0)} - \alpha_4 m_0 c \right] b_k^{(0)} = 0, \\ (b) \quad \left[\pi_4^{(0)} + \sum_{i=1}^4 \alpha_i \pi_i^{(0)} - \alpha_4 m_0 c \right] b_k^{(1)} = - \left(\frac{1}{c} \frac{\partial}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial}{\partial x_i} \right) b_k^{(0)}. \end{cases}$$

We then recover the same equations for the $b_k^{(i)}$ that we found for the $b_k^{(i)}$ in the Pauli argument.

Equations (X.h)-(a) admit solutions that are not identically zero only if their determinant is zero, which is verified because that determinant is equal to $(\pi_4^0)^2 - (\boldsymbol{\pi}^{(0)})^2 - m_0^2 c^2$. We then get the two independent solutions:

$$A_1 = -\frac{\pi_z^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad A_2 = -\frac{\pi_x^{(0)} + i\pi_y^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad A_3 = 1, \quad A_4 = 0,$$

and

$$B_1 = -\frac{\pi_x^{(0)} + i\pi_y^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad B_2 = \frac{\pi_z^{(0)}}{\pi_4^{(0)} + m_0 c}, \quad B_3 = 0, \quad B_4 = 1.$$

However, it is essential to remark that although the functions $A_k(\pi_4^0, \boldsymbol{\pi}^{(0)})$ and $B_k(\pi_4^0, \boldsymbol{\pi}^{(0)})$ have the same forms that they have in Pauli's argument, they do not vary with x, y, z, t in the same fashion. Indeed, here we have:

$$\pi_4^0 = \frac{1}{c} \frac{\partial S'_0}{\partial t} - \varepsilon V \quad \text{and} \quad \boldsymbol{\pi}^{(0)} = -\text{grad } S'_0 - \frac{\varepsilon}{c} \mathbf{A},$$

and not:

$$\pi_4^0 = \frac{1}{c} \frac{\partial S_0}{\partial t} - \varepsilon V \quad \text{and} \quad \boldsymbol{\pi}^{(0)} = -\text{grad } S_0 - \frac{\varepsilon}{c} \mathbf{A},$$

and the functions S_0 and S'_0 are defined differently. In particular, in the case of a permanent field in which the energy keeps a constant value W , the quantity $c\pi_4^0$ will be given, not by $W - \varepsilon V(x, y, z)$, as in the Pauli case, but by the formula:

$$c\pi_4^0 = W - \varepsilon V(x, y, z) - \eta^2 U(x, y, z),$$

in such a way that here *the existence of proper moments is already manifesting itself in the zero-order approximation.*

Now, if we again set:

$$b_k'^{(0)} = C_1(x, y, z, t) A_k(\pi_4^0, \boldsymbol{\pi}^{(0)}) + C_2(x, y, z, t) B_k(\pi_4^0, \boldsymbol{\pi}^{(0)}),$$

as in Pauli's argument, then we will recover the *Pauli conditions*:

$$\sum_{k=1}^4 A_k^* \left(\frac{1}{c} \frac{\partial b_k'^{(0)}}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial b_k'^{(0)}}{\partial x_i} \right) = 0, \quad \sum_{k=1}^4 B_k^* \left(\frac{1}{c} \frac{\partial b_k'^{(0)}}{\partial t} - \sum_{i=1}^3 \alpha_i \frac{\partial b_k'^{(0)}}{\partial x_i} \right) = 0,$$

whose explicit forms we gave above.

Pauli's calculations show that here again we will have the continuity equation:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0,$$

if we set:

$$\rho = \sum_{k=1}^4 b_k'^{(0)*} b_k'^{(0)}, \quad \mathbf{f} = \rho \mathbf{u} = - \sum_{k=1}^4 b_k'^{(0)*} \boldsymbol{\alpha} c b_k'^{(0)}$$

in the zero-order approximation, and we will also find that:

$$\sum_{k=1}^4 b_k'^{(0)*} \alpha_4 b_k'^{(0)} = \rho \sqrt{1 - \beta^2}.$$

The last formula justifies *a posteriori* the relation that we have assumed in order to establish the Jacobi equation, while the validity of the equation of continuity with the definition of ρ and \mathbf{f} above seems to indicate that motion takes place in accord with the laws of Weysenhoff's point mechanics; at least, in the zero-order approximation. Since that kind of point mechanics contains the action of the electromagnetic fields on the proper moments, we are in complete contradiction to Pauli's conclusions. The question that one must pose is then that of knowing whether the argument that we just gave or Pauli's argument is the one that leads to an exact conclusion in regard to this issue. In order to resolve that difficult question, one can first try to carry out the complete calculations in some well-defined cases. Unfortunately, those calculations are generally extremely difficult. In order to orient ourselves, we shall perform the calculations in a

very simple case, which will already reveal a certain number of interesting aspects of the situation to us.

6. CALCULATING THE FUNCTIONS C_1 AND C_2 IN THE CASE OF AN ELECTRON MOVING IN THE DIRECTION OF A PERMANENT, HOMOGENEOUS MAGNETIC FIELD

Take the direction of the constant field \mathbf{H} to be the z -axis and suppose that the motion of the electron takes place along that axis.

We perform the calculations in parallel under the Pauli hypothesis (point mechanics without spin) and Weysenhoff's point mechanics.

The Pauli conditions that determine C_1 and C_2 are the following ones:

$$2 \left[\pi_4^{(0)} \frac{1}{c} \frac{\partial C_1}{\partial t} + \pi_x^{(0)} \frac{\partial C_1}{\partial x} + \pi_y^{(0)} \frac{\partial C_1}{\partial y} + \pi_z^{(0)} \frac{\partial C_1}{\partial z} \right] + \pi_z^{(0)} \left(\frac{\partial C_2}{\partial x} - i \frac{\partial C_2}{\partial y} \right) - (\pi_x^{(0)} - i \pi_y^{(0)}) \frac{\partial C_2}{\partial x} \\ = - C_1 \operatorname{div} \boldsymbol{\pi}^{(0)} + i C_1 (\operatorname{rot} \boldsymbol{\pi}^{(0)})_0 - C_1 [(\operatorname{rot} \boldsymbol{\pi}^{(0)})_x + i (\operatorname{rot} \boldsymbol{\pi}^{(0)})_y].$$

Now, here one can set:

$$A_x = -\frac{H}{2}y, \quad A_y = \frac{H}{2}x, \quad A_z = 0,$$

which will give:

$$H_x = 0, \quad H_y = 0, \quad H_z = H.$$

The Lagrange momenta have the values:

$$p_x = \pi_x + \frac{\varepsilon}{c} A_x = \pi_x - \frac{\varepsilon H}{c} y, \quad p_y = \pi_y + \frac{\varepsilon}{c} A_y = \pi_y + \frac{\varepsilon H}{c} x, \quad p_z = \pi_z.$$

In Pauli's theory:

$$\pi_x = \pi_x^{(0)}, \dots$$

With Weysenhoff's point mechanics, one will have:

$$\boldsymbol{\pi} = \boldsymbol{\pi}^{(0)} + \frac{U_0}{c} \frac{\mathbf{v}}{\sqrt{1-\beta^2}},$$

and $\boldsymbol{\pi}$ will differ from $\boldsymbol{\pi}^{(0)}$ only by a term of order U_0 , which we neglect, because in the final result, it will give us only a supplementary term in U_0^2 , and in the point mechanics that we have adopted, we neglect terms in U_0^2 . We shall then make coincide $\boldsymbol{\pi}$ with $\boldsymbol{\pi}^{(0)}$, where $\boldsymbol{\pi}^{(0)}$ can have some values that are different in Pauli's theory from their values in the other one, moreover.

Now, from Jacobi's theory, \mathbf{p} is the gradient (with the sign changed) of the Jacobi function (S_0 , for Pauli, S , for us), hence:

$$\text{rot } \mathbf{p} = 0.$$

One then infers that:

$$\frac{\partial \pi_x^{(0)}}{\partial z} = \frac{\partial \pi_z^{(0)}}{\partial x}, \quad \frac{\partial \pi_y^{(0)}}{\partial z} = \frac{\partial \pi_x^{(0)}}{\partial y}, \quad \frac{\partial \pi_x^{(0)}}{\partial y} = \frac{\partial \pi_y^{(0)}}{\partial x} + \frac{\varepsilon}{c} H.$$

By hypothesis, the electron describes the z -axis, so the trajectories of the same "class" will be circular helices around Oz . The projection of $\boldsymbol{\pi}^{(0)}$ onto the plane xOy has the length $(\varepsilon / 2c) Hr$, and it will be normal to the radius vector \mathbf{r} in that plane. One will then have:

$$\pi_x^{(0)} = \frac{\varepsilon}{2c} Hr \frac{y}{r} = \frac{\varepsilon}{2c} Hy, \quad \pi_y^{(0)} = \frac{\varepsilon}{2c} Hr \left(-\frac{x}{r} \right) = -\frac{\varepsilon}{2c} Hx,$$

so:

$$\frac{\partial \pi_y^{(0)}}{\partial x} - \frac{\partial \pi_x^{(0)}}{\partial y} = -\frac{\varepsilon}{c} H, \quad \frac{\partial \pi_y^{(0)}}{\partial x} + \frac{\partial \pi_x^{(0)}}{\partial y} = 0, \quad \frac{\partial \pi_x^{(0)}}{\partial x} + \frac{\partial \pi_y^{(0)}}{\partial y} = 0.$$

Since $d\rho_z / dt = v_z \partial \rho_z / \partial t = 0$, one will also have $\partial \pi_z^{(0)} / \partial z = 0$, and finally:

$$\text{div } \boldsymbol{\pi}^{(0)} = 0.$$

Briefly, in the cases of $C_1 = 0$ or $C_2 = 0$, the Pauli equations reduce to:

$$2 \left[\pi_4^{(0)} \frac{1}{c} \frac{\partial C_2}{\partial t} + \pi_x^{(0)} \frac{\partial C_2}{\partial x} + \pi_y^{(0)} \frac{\partial C_2}{\partial y} + \pi_z^{(0)} \frac{\partial C_2}{\partial z} \right] = +i \frac{\varepsilon}{c} H C_2, \quad C_1 = 0,$$

$$2 \left[\pi_4^{(0)} \frac{1}{c} \frac{\partial C_1}{\partial t} + \pi_x^{(0)} \frac{\partial C_1}{\partial x} + \pi_y^{(0)} \frac{\partial C_1}{\partial y} + \pi_z^{(0)} \frac{\partial C_1}{\partial z} \right] = -i \frac{\varepsilon}{c} H C_1, \quad C_2 = 0,$$

respectively. One finds that:

$$\frac{4\pi}{h} m_0 U_0 = -\frac{\varepsilon}{c} H.$$

Now, from formulas (VI.d), one will have:

$$\mathfrak{M}_z = \int_V \mu_z d\tau = \int_V \frac{\varepsilon h}{4\pi m_0 c} (|C_1|^2 - |C_2|^2) \frac{2m_0 c}{\Delta} d\tau,$$

with $\Delta = \pi_4^{(0)} + m_0 c$, for uniform, rectilinear motion of the particle along Oz , which will give:

$$\mathfrak{M}_z = \frac{\varepsilon h}{4\pi m_0 c} |C_1|^2 \frac{2m_0 c}{\Delta} \mathcal{V}$$

here. Now, C_1 is normalized by the relation (VI.c):

$$\int_{\mathcal{V}} \rho d\tau = \int_{\mathcal{V}} (|C_1|^2 + |C_2|^2) \frac{2\pi_4^{(0)}}{\Delta} d\tau = |C_1|^2 \frac{2\pi_4^{(0)}}{\Delta} \mathcal{V} = 1,$$

so:

$$\mathfrak{M}_z = \frac{\varepsilon h}{4\pi m_0 c} \frac{m_0 c}{\pi_4^{(0)}}.$$

The relation:

$$\frac{4\pi}{h} \eta U m_0 = -H \quad (\text{with } U = -\mathfrak{M}_z \cdot H)$$

gives:

$$\mathfrak{M}_z = \frac{\varepsilon h}{4\pi m_0 c} \sqrt{1 - \beta^2},$$

and there is indeed agreement.

We have then obtained the solution $C_1 = D_1 e^{\frac{2\pi i}{h} \int U d\tau}$, $C_2 = 0$ that corresponds to the orientation parallel to the field \mathbf{H} and the magnetic moment \mathfrak{M} (whose components \mathfrak{M}_x and \mathfrak{M}_y are zero, as one easily verifies).

One similarly finds the solution $C_1 = 0$, $C_2 = D_2 e^{\frac{2\pi i}{h} \int U d\tau}$, which corresponds to an anti-parallel orientation of \mathbf{H} and \mathfrak{M} with $\mathfrak{M}_z = -\frac{\varepsilon h}{4\pi m_0 c}$. D_1 and D_2 will be constants if \mathbf{H} is constant in space; they will vary slowly if \mathbf{H} is slowly-varying in space at the large scale.

The solutions obtained are rigorous in Pauli's theory. They are exact, up to terms in U^2 , in our theory, but those terms are assumed to be negligible.

We remark that our result proves that even in Pauli's theory, the action of the electromagnetic field on the proper moments will enter into the expression for the functions C_1 and C_2 , and therefore into those of the $b_k^{(0)}$; i.e., in the zero-order approximation. That amounts to saying that in Dirac's theory, when one introduces electromagnetic potentials into the Dirac equations, one will introduce the action of the field on the proper moments in so doing. One will likewise see that action appear when one takes the Jacobi function to be the function S_0 for the electron without spin. However, the terms in U that appear in the expression for an imaginary exponential will be eliminated when one forms the quadratic quantities that have physical significance, such as ρ and \mathbf{f} , and from that viewpoint, the Pauli argument does not seem to have been weakened.

In Pauli's theory, the result obtained will give:

$$\psi_k = C_k e^{\frac{2\pi i}{h}(S_0 + \int U d\tau)},$$

and in ours:

$$\psi_k = C_k e^{\frac{2\pi i}{h}(S'_0 + \int U d\tau)} = e^{\frac{2\pi i}{h}S},$$

but the velocity $v = \beta c$ that figures in the phase function will not be the same. The phase function can be written:

$$\int \left\{ \left[\frac{m_0 c^2}{\sqrt{1-\beta^2}} + \frac{U_0}{c\sqrt{1-\beta^2}} \right] dt - \sum_{i=1}^3 \left[\frac{m_0 v_i}{\sqrt{1-\beta^2}} + \frac{\epsilon}{c} A_i + \frac{U_0}{c^2} \cdot \frac{v_i}{\sqrt{1-\beta^2}} \right] dx_i \right\}$$

in both cases, but in the Pauli case $W = \frac{m_0 c^2}{\sqrt{1-\beta^2}}$, while in our case, one will have:

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + \frac{U_0}{c\sqrt{1-\beta^2}} = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + \frac{\epsilon h}{4\pi m_0 c} \cdot H \cdot \frac{1}{\sqrt{1-\beta^2}}.$$

It would seem that this represents a first advantage of our viewpoint over that of Pauli. Indeed, Weyssenhoff's point mechanics can make the velocity \mathbf{v} that it assumes correspond quite naturally to the Jacobi function $S = S'_0 + \int U d\tau$, as one will see upon referring to the formulas that we have developed. One can once more say that the velocity \mathbf{v} is the "group velocity" of that corresponds to the phase that is defined by S . On the contrary, in the Pauli conception of things, the velocity \mathbf{v} that one adopts corresponds to the Jacobi function S_0 of the particle without spin, and not to the function $S_0 + \int U d\tau$ that appears in the exponent of the exponential in the ψ_k ; therefore, that velocity \mathbf{v} is not equal to the "group velocity" that corresponds to the complex phase $S_0 + \int U d\tau$ of the ψ_k .

It seems that Pauli's mechanics and our own differ profoundly from the physical standpoint. Suppose that the field H is zero along the axis Oz in the region (1), which extends from $z = -\infty$ to $z = a$. It will then increase from 0 to H very slowly at the large scale in the region (2), which goes from $z = a$ to $z = b$. Finally, the magnetic field has the constant value H in the region (3); viz., for $z > b$. In the first region, an electron arrives along the Oz axis that is represented by a small train of monochromatic waves whose magnetic moment is either parallel or anti-parallel to Oz . The energy W of the electron will remain constant when it crosses the region (2) and arrives in the region (3). From Pauli's theory, its velocity (i.e., the velocity of the wave train) will also remain the same, since the action of the field on the proper moment must not be felt in the point mechanics approximation. From our viewpoint, which adopts Weyssenhoff's point mechanics, the collective velocity of the wave train will vary slowly when it traverses region (2), and will take on a constant value in region (3) that is different from its initial value, so that final value will be smaller or greater than the initial value according to whether the

magnetic moment of the electron is anti-parallel or parallel to the field, respectively. The predictions of the two theories then seem to be physically different.

7. RETURN TO THE PASSAGE TO GEOMETRICAL OPTICS IN DIRAC'S THEORY

When we studied the passage to the geometrical optics approximation in Dirac's theory in paragraph 3 of this chapter, we found the *rigorous* equation:

$$\pi_4^2 - \pi^2 + F(a_k) + G(a_k) = +2m_0 \eta U,$$

in which the a_k are the amplitudes of the ψ_k :

$$\psi_k = a_k e^{\frac{2\pi i}{h} S}.$$

The functions $F(a_k)$ and $G(a_k)$ have the expressions:

$$F(a_k) = \frac{h}{2\pi i} \frac{\int \sum_{j=1}^4 \sum_{k=1}^4 \left(a_k^* \alpha_4 \frac{\partial a_k}{\partial x_j} - \frac{\partial a_k^*}{\partial x_j} \alpha_4 a_k \right) \pi_j d\tau}{\sqrt{1-\beta^2}},$$

$$G(a_k) = -\frac{h^2}{8\pi^2} \frac{\int \sum_{k=1}^4 \left(a_k^* \alpha_4 \square a_k - \square a_k^* \alpha_4 a_k \right) d\tau}{\sqrt{1-\beta^2}}$$

(in which $x_4 = ct, \dots$).

The term $G(a_k)$, which has order h^2 , is certainly negligible when the conditions of geometrical optics are verified. The same thing will be true for $F(a_k)$, which has the expression:

$$F(a_k^{(0)}) = \frac{h}{2\pi i} \frac{\int \sum_{j=1}^4 \sum_{k=1}^4 \left(a_k^{(0)*} \alpha_4 \frac{\partial a_k^{(0)}}{\partial x_j} - \frac{\partial a_k^{(0)*}}{\partial x_j} \alpha_4 a_k^{(0)} \right) \pi_j^{(0)} d\tau}{\sqrt{1-\beta^2}}$$

to the zero-order of approximation, in the point mechanics that have adopted. Suppose that we have written the development of the ψ_k in the form:

$$\psi_k = \left(b_k^{(0)} + \frac{h}{2\pi i} b_k^{(1)} + \dots \right) e^{\frac{2\pi i}{h} S_0}.$$

Upon neglecting the terms in h^2 , we will get:

$$(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 + F(b_k'^{(0)}) = m_0^2 c^2 + 2m_0 \eta U.$$

$F(b_k'^{(0)})$ has the same expression as $F(a_k^{(0)})$, when one substitutes the $b_k'^{(0)}$ for the $a_k^{(0)}$. If the $F(b_k'^{(0)})$ are negligible then we will get:

$$(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 = m_0^2 c^2 + 2m_0 \eta U,$$

which is a relation that will be inexact, even when one supposes that K is negligible. In order to get an exact relation, one must have:

$$(X.i) \quad F(b_k'^{(0)}) = 2m_0 \eta U = 2m_0 U_0,$$

which will then indeed give the exact relation $(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 = m_0^2 c^2$. If one assumes the hypothesis that is expressed by the relation (X.i) then $F(a_k^{(0)})$ will be zero, because one easily proves that since:

$$b_k'^{(0)} = a_k^{(0)} e^{\frac{2\pi i}{h} \int U d\tau} + \text{terms of order } \frac{h}{2\pi i},$$

one will have:

$$F(b_k'^{(0)}) = F(a_k^{(0)}) + 2m_0 \eta U,$$

so $F(a_k^{(0)}) = 0$. The hypothesis that was made in the deduction of the Jacobi equation by neglecting the term $F(a_k^{(0)})$ is therefore verified. It then remains for us to examine whether the relation (X.i) is indeed exact.

Before we make that examination, we make the following remark: If, instead of reasoning as we just did – i.e., with Weyssenhoff's point mechanics – we take the Pauli viewpoint and its development $\psi_k = \left(b_k^{(0)} + \frac{h}{2\pi i} b_k^{(1)} + \dots \right) e^{\frac{2\pi i}{h} S_0}$, in which S_0 is the Jacobi function of the electron *without spin*, then the passage to the geometrical optics approximation will yield the equation (in the zero-order approximation):

$$(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 + F(b_k'^{(0)}) = m_0^2 c^2 + 2m_0 \eta U,$$

with:

$$\pi_4^{(0)} = \frac{m_0 c}{\sqrt{1 - \beta^2}}, \quad \boldsymbol{\pi}^{(0)} = \frac{m_0 \boldsymbol{\beta} c}{\sqrt{1 - \beta^2}}, \quad \eta = \frac{1}{\sqrt{1 - \beta^2}}.$$

Here, all other things being equal moreover, the velocity $\boldsymbol{\beta}c$ does not have the same value that it had in Weyssenhoff's point mechanics. Now, we obviously have that once more $(\pi_4^{(0)})^2 - (\boldsymbol{\pi}^{(0)})^2 = m_0^2 c^2$. It will then be necessary that one must have:

$$(X.j) \quad F(b_k^{(0)}) = 2m_0 \eta U,$$

in Pauli's theory, as well.

Briefly: Upon assuming the Pauli viewpoint, one must have equation (X.j), while upon assuming our viewpoint and Weysenhoff's point mechanics, one must have the relation (X.i) in order to be able to neglect $F(a_k^{(0)})$ in the Jacobi equation. The verification of relations (X.i) and (X.j) in the particular cases is difficult, due to the complexity in the calculations that one must perform. We shall indicate that fact in three simple special cases.

1. – *Longitudinal motion in a uniform magnetic field.*

This is a case that have studied previously. From our viewpoint, we found that:

$$\psi_k = C_k e^{\frac{2\pi i}{h} S_0'} e^{\frac{2\pi i}{h} \int U dt},$$

in the zero-order approximation and upon neglecting the terms in U^2 ; hence:

$$b_k^{(0)} = C_k e^{\frac{2\pi i}{h} \int \left[\frac{U_0}{\sqrt{1-\beta^2}} dt - \sum_{i=1}^3 \frac{U_0}{c^2 \sqrt{1-\beta^2}} v_i dx_i \right]},$$

with $C_k = A_k$ or B_k , according to whether the proper magnetic moment of the electron is parallel or anti-parallel to the field, resp., and A_k and B_k are suitably normalized.

Upon substituting that expression for the $b_k^{(0)}$ into $F(b_k^{(0)})$, one will easily find that:

$$F(b_k^{(0)}) = 2m_0 \eta U.$$

On the contrary, if one assumes the Pauli viewpoint then one must set:

$$\psi_k = C_k e^{\frac{2\pi i}{h} \int U dt} e^{\frac{2\pi i}{h} S_0},$$

in which S_0 is the Jacobi function for the particle without spin. One will then have:

$$b_k^{(0)} = C_k e^{\frac{2\pi i}{h} \int \left[\frac{U_0}{\sqrt{1-\beta^2}} dt - \sum_{i=1}^3 \frac{U_0}{c^2 \sqrt{1-\beta^2}} v_i dx_i \right]},$$

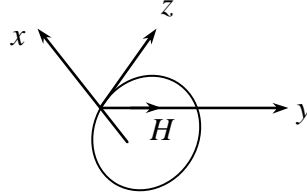
in which one always has $C_k = A_k$ or B_k ; however, β also corresponds to the velocity in point mechanics for the particle without spin. The substitution of $b_k^{(0)}$ into $F(b_k^{(0)})$ will then give:

$$F(b_k^{(0)}) = 2m_0 \eta U.$$

The relations (X.i) and (X.j) are then indeed verified in the present case.

2. – *Transversal motion in a uniform magnetic field.*

In a uniform magnetic field, consider a circular electron trajectory in a plane that is normal to the field. We shall calculate F at a point O on the trajectory. We take the tangent to the trajectory to be the z -axis, the direction of the uniform magnetic field to be the y -axis, and the normal to the trajectory to be the x -axis.



We begin here by performing the calculation while adopting Pauli's viewpoint; i.e., the point mechanics of a particle without spin. The equations of the Pauli condition take the simple form:

$$\frac{\partial C_1}{\partial z} = \frac{1}{2\pi_z^{(0)}} \frac{\varepsilon}{c} H C_2 - \frac{\pi_4^{(0)}}{\pi_z^{(0)}} \frac{1}{c} \frac{\partial C_1}{\partial t} - C_1 \operatorname{div} \boldsymbol{\pi},$$

$$\frac{\partial C_2}{\partial z} = -\frac{1}{2\pi_z^{(0)}} \frac{\varepsilon}{c} H C_1 - \frac{\pi_4^{(0)}}{\pi_z^{(0)}} \frac{1}{c} \frac{\partial C_2}{\partial t} - C_2 \operatorname{div} \boldsymbol{\pi}.$$

(The class of motions considered corresponds to the planar circular trajectories whose axes are along the field \mathbf{H} .) We now perform the calculation at the point O where:

$$\pi_x^{(0)} = 0, \quad \pi_y^{(0)} = 0,$$

and where:

$$\frac{\partial \pi_z^{(0)}}{\partial z} = 0 \quad \left(\text{because } \frac{d\pi_x^{(0)}}{dz} = 0 = v_z \frac{\partial \pi_x^{(0)}}{\partial z} \right).$$

We would now like to calculate the expression:

$$F(b_k^{(0)}) = \frac{h\eta}{2\pi i} \int d\tau \sum_{j=1}^4 \sum_{k=1}^4 \left(b_k^{(0)*} \alpha_4 \frac{\partial b_k^{(0)}}{\partial x_j} - \frac{\partial b_k^{(0)*}}{\partial x_j} \alpha_4 b_k^{(0)} \right) \pi_j^{(0)}.$$

It is easy to verify that the terms in that expression that contain the derivatives $\frac{1}{c} \frac{\partial}{\partial t}$ will cancel the terms in $\operatorname{div} \boldsymbol{\pi}^{(0)}$; they are zero, moreover, as we will see below. In order to calculate $F(b_k^{(0)})$ at 0, it will suffice to set:

$$\frac{\partial C_1}{\partial z} = \frac{1}{2\pi_z^{(0)}} \frac{\varepsilon}{c} H C_2, \quad \frac{\partial C_2}{\partial z} = -\frac{1}{2\pi_z^{(0)}} \frac{\varepsilon}{c} H C_1,$$

$\pi_y^{(0)} = 0$ or the motion considered, and one will have (with $\Delta = \pi_4^{(0)} + m_0 c$):

$$b_1^{(0)} = \frac{1}{\Delta}[-\pi_y^{(0)} C_1 - \pi_x^{(0)} C_2], \quad b_2^{(0)} = \frac{1}{\Delta}[\pi_z^{(0)} C_2 - \pi_y^{(0)} C_1], \quad b_3^{(0)} = C_1, \quad b_4^{(0)} = C_2,$$

so:

$$(X.k) \quad \begin{cases} \frac{\partial b_1^{(0)}}{\partial z} = \frac{1}{\Delta} \left[-C_2 \frac{\partial \pi_x^{(0)}}{\partial z} - \pi_z^{(0)} \frac{\partial C_2}{\partial z} \right] = \frac{1}{\Delta} \frac{C_2}{2} \frac{\varepsilon}{c} H, \\ \frac{\partial b_2^{(0)}}{\partial z} = \frac{1}{\Delta} \left[-C_1 \frac{\partial \pi_x^{(0)}}{\partial z} - \pi_z^{(0)} \frac{\partial C_1}{\partial z} \right] = \frac{1}{\Delta} \frac{C_1}{2} \frac{\varepsilon}{c} H, \end{cases}$$

because one has:

$$\pi_x^{(0)} = 0, \quad \frac{\partial \pi_z^{(0)}}{\partial z} = 0 \quad \text{and} \quad \frac{d}{dt} \pi_x^{(0)} = v_z \frac{\partial \pi_x^{(0)}}{\partial z} = -\frac{\varepsilon}{c} v_z H$$

at 0, so:

$$\frac{\partial \pi_x^{(0)}}{\partial z} = -\frac{\varepsilon}{c} H.$$

One then finds:

$$F(b_k^{(0)}) = \frac{-h\eta}{2\pi i} \int \pi_j^{(0)} d\tau \left\{ \left[\left(b_1^{(0)*} \frac{1}{\Delta} \frac{C_2}{2} + b_1^{(0)*} \frac{1}{\Delta} \frac{C_1}{2} \right) \frac{\varepsilon}{c} H - b_3^{(0)*} \frac{\partial C_1}{\partial z} - b_4^{(0)*} \frac{\partial C_2}{\partial z} \right] + \text{identical term} \right\}.$$

Hence:

$$(X.l) \quad \begin{aligned} F(b_k^{(0)}) &= \frac{-h\eta}{2\pi i} \int \pi_j^{(0)} d\tau \pi_z^{(0)} \left[-\frac{\pi_z^{(0)}}{\Delta^2} C_1^* C_2 + \frac{\pi_z^{(0)}}{\Delta^2} C_2^* C_1 - \frac{C_1^* C_2}{\pi_z^{(0)}} - \frac{C_2^* C_1}{\pi_z^{(0)}} \right] \frac{\varepsilon}{c} H. \\ &= -\frac{h\eta}{2\pi i} i \int \pi_j^{(0)} d\tau (C_1^* C_2 - C_2^* C_1) \left(1 + \frac{\pi_z^{(0)}}{\Delta^2} \right) \frac{\varepsilon}{c} H. \end{aligned}$$

Now:

$$\frac{1 + \pi_z^{(0)}}{\Delta^2} = \frac{(\pi_4^{(0)})^2 - m_0^2 c^2}{(\pi_4^{(0)} - m_0 c)^2} + 1 = \frac{2\pi_4^{(0)}}{\Delta},$$

so:

$$F(b_k^{(0)}) = -2m_0 \eta H i \int \frac{\varepsilon h}{4\pi m_0 c} (C_1^* C_2 - C_2^* C_1) \frac{2\pi_4^{(0)}}{\Delta} d\tau.$$

Formulas (VI.d) give:

$$\mu_y = i (C_1^* C_2 - C_2^* C_1) \frac{\varepsilon h}{4\pi m_0 c} \cdot \frac{2\pi_4^{(0)}}{\Delta} \quad (\text{with } \mathfrak{M}_y = \int \mu_y d\tau).$$

One then gets:

$$F(b_k^{(0)}) = -2m_0\eta \mathfrak{M}_y H,$$

and since $U = -\mathfrak{M}_y H$, one will have:

$$F(b_k^{(0)}) = 2m_0\eta U = 2m_0U_0.$$

The relation (X.j) is in fact verified.

Now, if we repeat the calculation from our own viewpoint – i.e., upon adopting Weyssenhoff's point mechanics – then the argument will stay the same, except that here we will have to set:

$$\mathbf{p} = \boldsymbol{\pi}^{(0)} \cdot \frac{U_0\eta}{c^2} \mathbf{v} \quad \text{and} \quad \frac{d\mathbf{p}}{dt} = \mathbf{f}.$$

We will then infer that:

$$v_z \frac{\partial}{\partial z} \left(\pi_x^{(0)} + \frac{U_0\eta}{c^2} v_x \right) = -\frac{\varepsilon}{c} v_z H,$$

so once more:

$$\frac{\partial \pi_z^{(0)}}{\partial z} = 0,$$

because v_z will remain constant to first order on the trajectory. We will then have:

$$v_z \frac{\partial}{\partial z} \left(\pi_x^{(0)} + \frac{U_0\eta}{c^2} v_x \right) = -\frac{\varepsilon}{c} v_z H,$$

so:

$$\frac{\partial \pi_x^{(0)}}{\partial z} = -\frac{\varepsilon}{c} H - \frac{\partial}{\partial z} \left(\frac{U_0\eta v_x}{c^2} \right),$$

instead of:

$$\frac{\partial \pi_x^{(0)}}{\partial z} = -\frac{\varepsilon H}{c}.$$

That will introduce some terms in U into the last side of equations (X.j), in addition to the terms in $\varepsilon/c H$. We will then obtain an expression for $F(b_k'^{(0)})$ that has the same form as the one that was given above for $F(b_k^{(0)})$, but in which the factor $\varepsilon/c H$ is replaced with $\varepsilon/c H +$ a term in U . In the expression for $F(b_k'^{(0)})$, those terms will then give us some supplementary terms in h^2 that are negligible. We will then have:

$$F(b_k'^{(0)}) = 2m_0 \eta U,$$

and the relation (X.i) will be verified. [More precisely, we find that $F(b_k^{(0)}) = 2m_0 \eta U \left(1 + \frac{U}{m_0 c^2}\right)$, and we neglect terms in U^2 .]

One can easily integrate the differential equations in C_1 and C_2 that were given at the beginning of the study of this second special case. One has $\frac{\partial \pi_z^{(0)}}{\partial z} = 0$, $\frac{\partial \pi_y^{(0)}}{\partial y} = 0$, $\frac{\partial \pi_x^{(0)}}{\partial x} = 0$ for the class of motions that includes all planar circular trajectories whose axis is along \mathbf{H} .

The equations in C_1 and C_2 can then be written (since $\boldsymbol{\pi}^{(0)} = 0$):

$$m_0 \eta \frac{\partial C_1}{\partial t} = \frac{1}{2} \frac{\varepsilon}{c} H C_2, \quad m_0 \eta \frac{\partial C_2}{\partial t} = -\frac{1}{2} \frac{\varepsilon}{c} H C_1,$$

and one will find the solution:

$$C_1 = D e^{\frac{2\pi i}{h} \int U dt \sqrt{1-\beta^2}}, \quad C_2 = \pm i C_1.$$

One can take the + or – sign in the second equation according to whether the proper magnetic moment is parallel to anti-parallel to \mathbf{H} , resp.

It might then seem that upon introducing the preceding expressions into that of $F(b_k^{(0)})$ [or $F(b_k^{\prime(0)})$], one must obtain:

$$\frac{F(b_k^{(0)})}{F(b_k^{\prime(0)})} = 2m_0 \eta U \sqrt{1-\beta^2} = 2m_0 U,$$

contrary to the result above, but that is not true, because one must take into account some terms in $\partial \pi_x^{(0)} / \partial z$ that express the variation of $\pi_x^{(0)}$. In order to better account for what happens, recall formula (X.l) and describe it in the form:

$$F(b_k^{(0)}) = -\frac{\hbar \eta}{2\pi} i \int d\tau (C_1^* C_2 - C_2^* C_1) \left(1 - \frac{(\pi_z^{(0)})^2}{\Delta^2} + 2 \frac{(\pi_x^{(0)})^2}{\Delta^2} \right) \frac{\varepsilon}{c} H.$$

The term in $1 - (\pi_x^{(0)})^2 / \Delta^2$ is the one that one will obtain if $\partial \pi_x^{(0)} / \partial z$ is zero, as one sees upon repeating the calculations that were performed before in this special case. It will then be the one that is due to the variation of the C_i , while the term in $2(\pi_x^{(0)})^2 / \Delta^2$ is provided by the variation of the A_k and B_k .

As one must expect, the first term gives:

$$2m_0 \eta U \frac{1 - (\pi_x^{(0)})^2}{2\pi_x^{(0)} / \Delta} = 2m_0 \eta U \frac{2m_0 c / \Delta}{2\pi_x^{(0)} / \Delta} = 2m_0 \eta U \sqrt{1-\beta^2} = 2m_0 U,$$

but the second one gives:

$$2m_0 \eta U \frac{2(\pi_z^{(0)})^2}{\Delta^2} \cdot \frac{\Delta}{2\pi_4^{(0)}} = 2m_0 U \frac{\eta(\pi_z^{(0)})^2}{\Delta\pi_4^{(0)}},$$

in such a way that in total, one will have:

$$F(b_k^{(0)}) = 2m_0 U \left(1 + \eta \frac{(\pi_z^{(0)})^2}{\Delta\pi_4^{(0)}} \right).$$

Now:

$$1 + \eta \frac{(\pi_z^{(0)})^2}{\Delta\pi_4^{(0)}} = 1 + \eta \frac{(\pi_z^{(0)})^2 - m_0^2 c^2}{(\pi_4^{(0)})^2 + m_0 c \pi_4^{(0)}} = 1 + \eta \frac{m_0^2 c^2 (\eta^2 - 1)}{m_0^2 c^2 \eta(\eta + 1)} = 1 + \frac{\eta^2 - 1}{\eta + 1} = \eta,$$

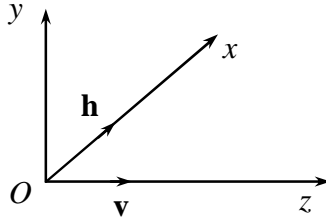
because $\eta_4^{(0)} = m_0 c \eta$, so finally:

$$F(b_k^{(0)}) = 2m_0 U.$$

[One will get the same result for $F(b_k^{(0)})$, up to terms in U^2 .]

3. – Transverse motion in a uniform electric field.

We shall suppose that the motion takes place in an electric field \mathbf{h} whose direction we take to be the Ox axis. We suppose that at the point O in question along the trajectory, the motion takes place along the Oz axis:



The trajectory will be a parabola in the plane xOz that is tangent to Oz and has its summit at O . We take the class of parabolic motions to be the ones that are obtained by displacing it parallel to itself along Oz and along Oy . In the plane yOz , the axis Oz will be a caustic, which will give rise to some singularities.

Here we have:

$$\pi_y^{(0)} = 0, \quad V = -h_x \cdot x, \quad -\frac{\partial V}{\partial x} = h_x = h,$$

$$b_1^{(0)} = -\frac{1}{\Delta}(\pi_z^{(0)}C_1 + \pi_x^{(0)}C_2), \quad b_2^{(0)} = -\frac{1}{\Delta}(\pi_z^{(0)}C_2 - \pi_x^{(0)}C_1), \quad b_3^{(0)} = C_1, \quad b_4^{(0)} = C_2.$$

One must take into account the variation of $\Delta = \pi_4^{(0)} + m_0 c$ that results from the action of the electric field on the particle. Here, the function S_0 is:

$$S_0 = Wt - \int \sqrt{\left(\frac{W}{c} + \frac{\varepsilon}{c} h_x x\right)^2 - (\pi_z^{(0)})^2 - m_0^2 c^2} dx - \int \pi_z^{(0)} dz,$$

with $\pi_z^{(0)} = \text{const.}$, and the integrals are taken along the trajectory. One finds that:

$$\pi_x^{(0)} = \sqrt{\left(\pi_4^{(0)} + \frac{\varepsilon}{c} h_x x\right)^2 - (\pi_z^{(0)})^2 - m_0^2 c^2}, \quad \text{so} \quad \frac{\partial \pi_x^{(0)}}{\partial x} = \frac{\pi_4^{(0)} \varepsilon}{\pi_x^{(0)} c} h_x.$$

$\partial \pi_x^{(0)} / \partial x$ is infinite at the origin due to the role of the caustic on the Oz axis (because $\pi_x^{(0)} = 0$ at the origin O).

Remark. – One can recover the preceding relations in the following way: On the parabolic trajectory, one will have:

$$z = v_z t, \quad \text{with} \quad v_z = \text{const.}, \quad x = \frac{1}{2} \frac{\varepsilon h_x}{m} t^2, \quad \text{so} \quad v_x = \frac{dx}{dt} = \frac{\varepsilon h_x}{m} t.$$

One will then have:

$$z^2 = 2 \frac{mv_z^2}{\varepsilon h_x} x = 2px, \quad \text{with} \quad p = \frac{mv_z^2}{\varepsilon h_x},$$

so one will find that:

$$\frac{dz}{dx} = \frac{p}{z}.$$

One also finds that:

$$\pi_z^{(0)} = m_0 h v_z, \quad \text{so} \quad \frac{\partial \pi_z^{(0)}}{\partial z} = 0 \text{ at } 0,$$

and then:

$$\frac{d\pi_x^{(0)}}{dt} = \varepsilon h_x = v_z \frac{\partial \pi_x^{(0)}}{\partial z},$$

which gives:

$$\frac{\partial \pi_x^{(0)}}{\partial z} = \frac{\varepsilon h_x}{v_z}.$$

One infers from this that:

$$\frac{\partial \pi_x^{(0)}}{\partial x} = \frac{\partial \pi_x^{(0)}}{\partial z} \frac{dz}{dx} = \frac{\varepsilon h_x}{v_z} \frac{p}{z} = \frac{mv_z}{z} = \frac{m}{t}.$$

Now, at O :

$$\frac{\pi_4^{(0)} \varepsilon}{\pi_x^{(0)} c} h_x = \frac{m_0 \eta c}{m_0 \eta v_x} \frac{\varepsilon h_x}{c} = \frac{\varepsilon h_x}{v_x} = \frac{m}{t},$$

so:

$$\frac{\partial \pi_x^{(0)}}{\partial x} = \frac{\pi_4^{(0)}}{\pi_x^{(0)}} \frac{\varepsilon}{c} h_x.$$

This must be proved.

On the other hand:

$$\Delta = \pi_4^{(0)} + m_0 c = \left[\frac{W}{c} + \frac{\varepsilon}{c} h_x x \right],$$

so:

$$\frac{\partial \Delta}{\partial x} = \frac{\varepsilon}{c} h_x \quad \text{and} \quad \frac{\partial \Delta}{\partial z} = 0.$$

We now have to write the Pauli equations. They are not given by formulas (X.f) here, because since $\pi_4^{(0)}$ varies, one must take into account the derivatives of Δ . One finds that:

$$\begin{aligned} \frac{\partial C_1}{\partial z} &= \frac{C_1}{2\pi_z^{(0)}} \frac{\partial \pi_x^{(0)}}{\partial x} + \frac{1}{2\pi_z^{(0)} \Delta} (\pi_z^{(0)} C_2 - \pi_x^{(0)} C_1) \frac{\partial \Delta}{\partial x} - \frac{\pi_4^{(0)}}{\pi_z^{(0)}} \frac{1}{c} \frac{\partial C_1}{\partial t}, \\ \frac{\partial C_2}{\partial z} &= \frac{C_2}{2\pi_z^{(0)}} \frac{\partial \pi_x^{(0)}}{\partial x} + \frac{1}{2\pi_z^{(0)} \Delta} (\pi_z^{(0)} C_1 + \pi_x^{(0)} C_2) \frac{\partial \Delta}{\partial x} - \frac{\pi_4^{(0)}}{\pi_z^{(0)}} \frac{1}{c} \frac{\partial C_2}{\partial t}. \end{aligned}$$

Here again, the terms in $\frac{1}{c} \frac{\partial}{\partial t}$ will cancel when one forms $F(b_k^{(0)})$, and we will have to employ the preceding expressions for $\frac{\partial C_1}{\partial z}$ and $\frac{\partial C_2}{\partial z}$, while omitting the last terms. That will then give:

$$F(b_k^{(0)}) = \frac{h\eta}{2\pi i} \int \left[\pi_x^{(0)} \left(\sum_{k=1}^4 b_k^{(0)*} \alpha_4 \frac{\partial b_k^{(0)}}{\partial x} - \text{conj.} \right) + \pi_z^{(0)} \left(\sum_{k=1}^4 b_k^{(0)*} \alpha_4 \frac{\partial b_k^{(0)}}{\partial z} - \text{conj.} \right) \right] d\tau.$$

One might believe that the first term in the bracket is zero, since one has $\pi_x^{(0)} = 0$ at 0. However, that is not at all true, because $\partial \pi_x^{(0)} / \partial x$ is proportional to $1 / \pi_x^{(0)}$, in such a way that the product $\pi_x^{(0)} \cdot \partial \pi_x^{(0)} / \partial x$ will remain finite and equal to $\pi_x^{(0)} (\varepsilon / c) h_x$ when $\pi_x^{(0)}$ tends to zero. The terms in $\partial / \partial x$ will then give:

$$-2 (C_1^* C_2 - C_2^* C_1) \frac{(\pi_4^{(0)})^2}{\Delta^2} \frac{\varepsilon}{c} h_x,$$

while the terms in $\partial / \partial z$ will give:

$$\left(-\frac{1}{\Delta^2} (\pi_z^{(0)})^2 C_1^* \frac{\partial C_1}{\partial z} - \frac{1}{\Delta^2} (\pi_z^{(0)})^2 C_2^* \frac{\partial C_2}{\partial z} + C_1^* \frac{\partial C_1}{\partial z} + C_2^* \frac{\partial C_2}{\partial z} \right) \pi_z^{(0)}$$

when $\pi_x^{(0)}$ tends to zero, and when one takes the values of $\partial C_1/\partial z$ and $\partial C_2/\partial z$ into account, that will become:

$$-2(C_1^*C_2 - C_2^*C_1) \left(\frac{(\pi_4^{(0)})^2}{\Delta^2} - 1 \right) \frac{1}{\Delta} \frac{\partial \Delta}{\partial x} \pi_z^{(0)},$$

so

$$\begin{aligned} F(b_k^{(0)}) &= -\frac{\hbar\eta}{2\pi i} \int d\tau (C_1^*C_2 - C_2^*C_1) \left[\left(\frac{(\pi_4^{(0)})^2}{\Delta^2} - 1 \right) \frac{\pi_z^{(0)}}{\Delta} \frac{\partial \Delta}{\partial x} - \frac{2\pi_z^{(0)}\pi_4^{(0)}}{\Delta^2} \frac{\varepsilon}{c} h_x \right], \\ &= -\frac{\hbar\eta}{2\pi i} \int i(C_1^*C_2 - C_2^*C_1) \left[\left(\frac{(\pi_4^{(0)})^2}{\Delta^2} - 1 \right) \frac{\pi_z^{(0)}}{\Delta} + \frac{2\pi_z^{(0)}\pi_4^{(0)}}{\Delta^2} \right] \frac{\varepsilon}{c} h_x d\tau, \\ &= -2m_0\eta \int i(C_1^*C_2 - C_2^*C_1) \frac{\varepsilon h}{4\pi m_0 c} h_x \left(\frac{2m_0 c}{\Delta} \cdot \frac{\pi_z^{(0)}}{\Delta} + \frac{2\pi_z^{(0)}\pi_4^{(0)}}{\Delta^2} \right) d\tau, \\ &= -2m_0\eta h_x \int i(C_1^*C_2 - C_2^*C_1) \frac{\varepsilon h}{4\pi m_0 c} \cdot \frac{2\pi_z^{(0)}}{\Delta} \cdot d\tau. \end{aligned}$$

Now, one has [see formula (VI.d)]:

$$\mathfrak{P}_x = \int \pi_x d\tau = \int i(C_1^*C_2 - C_2^*C_1) \frac{\varepsilon h}{4\pi m_0 c} \cdot \frac{2\pi_z^{(0)}}{\Delta} \cdot d\tau,$$

so:

$$F(b_k^{(0)}) = -2m_0 \eta \mathfrak{P}_x h_x = 2m_0 \eta U = 2m_0 U_0,$$

which is again the relation (X.j). One will easily show that upon adopting Weysenhoff's point mechanics, and upon neglecting the terms in U^2 , one will likewise obtain the relation (X.i); i.e., $F(b_k^{(0)}) = 2m_0 \eta U$.

8. GENERAL PROOF OF FORMULAS (X.i) AND (X.j) UPON STARTING WITH THE GORDON DECOMPOSITION

We were able to verify the formulas (X.i) and (X.j) by studying some special cases. We shall now point out a general proof by appealing to the decomposition of the space-time vector of *density-current* that was given by Gordon and developed in paragraph 4 of Chapter VI. We shall write that decomposition while considering the density-current quadri-vector to describe the number of particles, and not the electric charge and electric current density, which amounts to dividing the expressions that were given previously at the indicated place by the electron charge $-e$. We shall distinguish covariant components from contravariant ones. We set:

$$f^1 = \sum_{k=1}^4 \psi_k^* (-c\alpha_1) \psi_k, \quad f^2 = \sum_{k=1}^4 \psi_k^* (-c\alpha_2) \psi_k, \quad f^3 = \sum_{k=1}^4 \psi_k^* (-c\alpha_3) \psi_k, \quad f^4 = c \sum_{k=1}^4 \psi_k^* \psi_k,$$

and

$$(X.m) \quad g^1 = -\frac{h}{4\pi i m_0} \sum_{k=1}^4 \left(\psi_k^* \alpha_4 \frac{\partial \psi_k}{\partial x} - \frac{\partial \psi_k^*}{\partial x} \alpha_4 \psi_k \right), \dots, \quad g^4 = \frac{h}{4\pi i m_0} \sum_{k=1}^4 \left(\psi_k^* \alpha_4 \frac{1}{c} \frac{\partial \psi_k}{\partial t} - \frac{1}{c} \frac{\partial \psi_k^*}{\partial t} \alpha_4 \psi_k \right),$$

which defines the space-time vector \mathbf{f} . We can then write the Gordon decomposition in the following form:

$$f^j = g^j + \frac{c}{\varepsilon} \sum_i \frac{\partial \mu^{ij}}{\partial x^i} \quad (j = 1, 2, 3, 4),$$

with:

$$\mu^{ij} = -\mu^{ji}: \quad \mu^{14} = \pi_x, \quad \mu^{24} = \pi_y, \quad \mu^{34} = \pi_z, \quad \mu^{23} = \mu_x, \quad \mu^{31} = \mu_y, \quad \mu^{12} = \mu_z,$$

in which $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ are the *proper electric moment density* and the *proper magnetic moment density*, respectively.

With Pauli, we set:

$$\psi_k = \left(b_k^{(0)} + \frac{h}{2mi} b_k^{(1)} + \dots \right) e^{\frac{2\pi i}{h} S_0}.$$

Let $g_{(S_0)}^j$ denote the part of the expression for g^j that contains the derivatives of S_0 and let $g_{(b)}^j$ denote the part of the expression for g^j that contains the derivatives of the \mathbf{b} . We will have:

$$g^j = g_{(S_0)}^j + g_{(b)}^j.$$

The $g_{(S_0)}^j$ have the same expressions as they do for a plane wave, so one easily verifies that:

$$f^j = g_{(S_0)}^j,$$

so:

$$g_{(b)}^j = -\frac{\varepsilon}{c} \sum_{i=1}^4 \frac{\partial \mu^{ij}}{\partial x^i}.$$

In the zero-order approximation – i.e., when one keeps only the $b_k^{(0)}$ – when one scalar multiplies \mathbf{g} by $\boldsymbol{\pi}^{(0)}$, and then multiplies the latter equation by $\pi_j^{(0)}$ and sums over j , one will get:

$$\sum_j g_{(b)}^j \pi_j^{(0)} = -\frac{\varepsilon}{c} \sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^i} \pi_j^{(0)}.$$

Now, upon comparing the expression (X.m) in the zero-order approximation with the expression for $F(b_k^{(0)})$, one will find that:

$$\int \sum_j g_{(b)}^j \pi_j^{(0)} d\tau = -\frac{1}{2m_0\eta} F(b_k^{(0)}),$$

so

$$F(b_k^{(0)}) = \frac{2m_0\eta c}{\varepsilon} \int \sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^j} \pi_j^{(0)} d\tau = \frac{2m_0c}{\varepsilon} \int \sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^j} \pi_j^{(0)} d\tau_0,$$

because:

$$d\tau = d\tau_0 \sqrt{1-\beta^2} = \frac{d\tau_0}{\eta}.$$

Upon developing Weysenhoff's theory, we found the relation (IV.h):

$$\sum_j \mu^{ij} \pi_j^{(0)} = 0,$$

so:

$$\sum_{i,j} \frac{\partial}{\partial x^i} (\mu^{ij} \pi_j^{(0)}) = 0, \quad \text{or} \quad \sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^i} \pi_j^{(0)} = -\sum_{i,j} \mu^{ij} \frac{\partial \pi_j^{(0)}}{\partial x^i}.$$

Upon noting that $\boldsymbol{\pi}^{(0)} = \text{grad } S_0 + (\varepsilon/c) \mathbf{A}$ and $\mu^{ij} = -\mu^{ji}$, we will find that:

$$\sum_{i,j} \mu^{ij} \frac{\partial \pi_j^{(0)}}{\partial x^i} = \sum_{i,j} \mu^{ij} \left(\frac{\partial \pi_j^{(0)}}{\partial x^i} - \frac{\partial \pi_i^{(0)}}{\partial x^j} \right) = \frac{\varepsilon}{2c} \sum_{i,j} \mu^{ij} F_{ij},$$

in which the F_{ij} are the components of the *electromagnetic field* tensor:

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}.$$

Hence:

$$\sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^i} \pi_j^{(0)} = -\frac{\varepsilon}{c} \sum_{i,j} \frac{1}{2} \mu^{ij} F_{ij}.$$

Now, from a relation in paragraph 3 of this chapter, we know that:

$$U_0 = \eta U = -\frac{1}{2} \int \sum_{i,j} \mu^{ij} F_{ij} d\tau_0,$$

so:

$$\int \sum_{i,j} \frac{\partial \mu^{ij}}{\partial x^i} \pi_j^{(0)} d\tau_0 = \frac{\varepsilon}{c} U_0.$$

We then obtain:

$$F(b_k^{(0)}) = 2m_0 \left(-\frac{1}{2} \int \sum_{i,j} \mu^{ij} F_{ij} d\tau_0 \right) = 2m_0 U_0 = 2m_0 \eta U,$$

If we repeat the same calculation upon starting with the development:

$$\psi_k = \left(b_k^{(0)} + \frac{\hbar}{2\pi i} b_k^{(1)} + \dots \right) e^{\frac{2\pi i}{\hbar} S_0}$$

then upon neglecting the terms in U^2 , we will find formula (X.i):

$$F(b_k^{(0)}) = 2m_0 \eta U.$$

9. INTERPRETATION OF THE CONDITION $F(a_k^{(0)}) = 0$.

In summary, from the preceding, the condition that characterizes the use of Weyssenhoff's point mechanics for the Dirac electron and which seems justified from the standpoint of wave mechanics is the condition that:

$$(X.n) \quad F(a_k^{(0)}) = 0,$$

in which the $a_k^{(0)}$ are defined by the development:

$$\psi_k = \left(a_k^{(0)} + \frac{\hbar}{2\pi i} a_k^{(1)} + \dots \right) e^{\frac{2\pi i}{\hbar} S},$$

in which S is the Jacobi function of Weyssenhoff point mechanics. What is the significance of that physical condition?

In order to recognize that significance, recall the formula:

$$f^j = g^j + \sum_i \frac{c}{\varepsilon} \frac{\partial \mu^{ij}}{\partial x^i}$$

that was written above and decompose the quadri-vector \mathbf{g} in the following fashion:

$$g^j = g_{(S)}^j + g_{(a)}^j,$$

in which $g_{(S)}^j$ is the part of g^j that is provided by the variation of S , and $g_{(a)}^j$ is the part that is provided by the variation of the a_k , which will reduce to $g_{(a_0)}^j$ in the zero-order approximation. One easily verifies then that the condition (X.n) is equivalent to the following one:

$$\int \sum_j g_{(a_0)}^j \pi_j^{(0)} d\tau = 0,$$

or, for a very small wave packet:

$$\sum_j g_{(a_0)}^j \pi_j^{(0)} = 0.$$

Since the $\pi_j^{(0)}$ are proportional to components u_j of the world-velocity of the particle, (since $\pi_j^{(0)} = m_0 c u_j$), one will also have:

$$\sum_j g_{(a_0)}^j u_j = 0.$$

Now, the left-hand side of that formula represents the scalar product in space-time of the quadri-vector $\mathbf{g}_{(a_0)}$ with the world-velocity \mathbf{u} . It will then be equal to the temporal component $(g_{(a_0)}^4)_0$ of $\mathbf{g}_{(a_0)}$ in the proper system of the particle.

Finally, we arrive at the following conclusion:

The temporal component of the quadri-vector \mathbf{g} in the proper system of the particle is the same as if the $a_k^{(0)}$ were constant.

In reality, the $a_k^{(0)}$ are not constant, in general, since they have the form:

$$a_k^{(0)} = b_k^{(0)} e^{-\frac{2\pi i}{h} \int U d\tau} = (C_1 A_k + C_2 B_k) e^{-\frac{2\pi i}{h} \int U d\tau},$$

and the factors $C_1, C_2, A_k, B_k, e^{-\frac{2\pi i}{h} \int U d\tau}$ are not, in general, constants, but in the expression for $(g_{(a_0)}^4)_0$, the variations of those various factors cancel each other. We shall verify that in the next paragraph by repeating the study of longitudinal and transversal motion in a uniform magnetic field in detail. One might remark that, in general, the spatial part of $\mathbf{g}_{(a_0)}$ in the proper system is not zero.

Now, return to the equation:

$$f^j = g^j + \frac{c}{\varepsilon} \sum_i \frac{\partial \mu^{ij}}{\partial x^i},$$

which we write:

$$\mathbf{f} = \mathbf{g} + \mathbf{h},$$

upon introducing a quadri-vector \mathbf{h} that is defined $h^j = \frac{c}{\varepsilon} \sum_i \frac{\partial \mu^{ij}}{\partial x^i}$. The quadri-vector \mathbf{f}

describes the total motion at the fine scale of the probability of presence of the particle, as it is imagined by point mechanics. \mathbf{f} is the sum of \mathbf{g} and \mathbf{h} , while the latter quadri-vector corresponds to a kind of internal motion or spin. One might see that by studying the

motion of Darwin's probability globule, for example (see the author's book *L'Electron Magnétique*, pp. 169.)

The component g^4 represents a sort of "mean" probability density that is obtained by neglecting the internal motion that is coupled to the spin, which is a motion that has no interest in point mechanics. As a result of the compensation for the effects of the electromagnetic field, when that probability density g^4 is evaluated in the proper system of the particle that is defined by Weysenhoff point mechanics, it will be found to be the same as it is for a probability globule in the absence of a field. That property seems to be characteristic of Weysenhoff's point mechanics, and that will justify its use for the Dirac electron.

10. VERIFICATION OF THE RELATION $(g_{(a_0)}^4) = 0$ IN TWO SIMPLE SPECIAL CASES

It is easy to verify the relation $(g_{(a_0)}^4) = 0$ in the two special simple cases of longitudinal and transversal motion in a uniform magnetic field that were studied previously.

In the case of a longitudinal motion (paragraph 6 of this chapter and the first example in paragraph 7 of the same chapter), one finds that either:

$$\psi_k = D A_k e^{\frac{2\pi i}{h} S} \quad \text{or} \quad \psi_k = D B_k e^{\frac{2\pi i}{h} S},$$

with D a constant, according to whether the magnetic moment is oriented parallel or anti-parallel to the magnetic field. The calculations are immediate, and give:

$$g_{(S)}^4 = \rho \frac{m'_0}{m_0}, \quad g_{(S)}^3 = \rho \beta c \frac{m'_0}{m_0}, \quad g_{(S)}^1 = g_{(S)}^2 = 0$$

$$g_{(a_0)}^1 = g_{(a_0)}^2 = g_{(a_0)}^3 = g_{(a_0)}^4 = 0.$$

The quadri-vector $\mathbf{g}_{(a_0)}$ is zero, here. One will then have the relation $g_{(a_0)}^4 = 0$ in the proper system.

The calculations are somewhat more complicated in the case of the transverse magnetic field (second example of paragraph 7 in that chapter). One will then have:

$$b_k^{(0)} = C_1 A_k + C_2 B_k,$$

with $C_2 = \pm i C_1$ and $C_1 = D e^{\frac{2\pi i}{h} \int U dt \sqrt{1-\beta^2}}$, $D = \text{const.}$, so:

$$a_k^{(0)} = D (A_k \pm i B_k) e^{\frac{2\pi i}{h} \int U dt (\sqrt{1-\beta^2} - 1)}.$$

Calculation again gives:

$$g_{(s)}^4 = \rho \frac{m'_0}{m_0}, \quad g_{(s)}^3 = \rho \beta c \frac{m'_0}{m_0}, \quad g_{(s)}^1 = g_{(s)}^2 = 0$$

and one will then find that:

$$\begin{aligned} g_{(a_0)}^4 &= \rho \frac{U_0}{m_0 c^2} \left[\sqrt{1-\beta^2} - 1 \right], \\ g_{(a_0)}^1 &= g_{(a_0)}^2 = 0, \\ g_{(a_0)}^3 &= \rho \beta c \frac{U_0}{m_0 c^2} \left[\sqrt{1-\beta^2} - 1 \right] - \frac{1}{\Delta} \beta U_0 \rho \\ &= \rho \beta c \frac{U_0}{m_0 c^2} \left[\left(\sqrt{1-\beta^2} - 1 \right) - \frac{1-\beta^2}{\sqrt{1-\beta^2} + 1} \right] \\ &= - \rho \beta c \frac{U_0}{m_0 c^2} \cdot \frac{1}{\sqrt{1-\beta^2} + 1}. \end{aligned}$$

The two terms in the first expression for $g_{(a_0)}^3$ correspond to the variations in the exponential that appears in the $a_k^{(0)}$ and the variations in the A_k and B_k , respectively.

One passes from the system of the observer (which is a well-defined Galilean system, since the electromagnetic field reduces to a magnetostatic field in it) to the proper system of the corpuscle by a Lorentz transformation that will give:

$$(g_{(a_0)}^4)_0 = \frac{g_{(a_0)}^4 - \frac{\beta}{c} g_{(a_0)}^3}{\sqrt{1-\beta^2}} = \rho \frac{U_0}{m_0 c^2} \left[\left(\sqrt{1-\beta^2} - 1 \right) + \frac{\beta^2}{\sqrt{1-\beta^2} + 1} \right] = 0,$$

$$(g_{(a_0)}^1)_0 = g_{(a_0)}^1 = 0, \quad (g_{(a_0)}^2)_0 = g_{(a_0)}^2 = 0,$$

$$(g_{(a_0)}^3)_0 = \frac{g_{(a_0)}^3 - \beta c g_{(a_0)}^4}{\sqrt{1-\beta^2}} = \rho \frac{U_0}{m_0 c^2} \frac{1}{\sqrt{1-\beta^2}} \left[-\frac{\beta^2}{\sqrt{1-\beta^2} + 1} - \beta c \left(\sqrt{1-\beta^2} - 1 \right) \right],$$

when it is applied to the components of the quadri-vector $\mathbf{g}_{(a_0)}$. The relation $(g_{(a_0)}^4)_0 = 0$ is therefore indeed verified.

One might be amazed to see β appear in the expression for $(g_{(a_0)}^3)_0$, which relates to the proper system, but one must note that the velocity βc has an absolute sense here, because it is the relative velocity of the proper system of the corpuscle with respect to the

system in which the electromagnetic field is purely magnetic. In a Galilean system that possesses an *arbitrary* relative velocity (in the sense of z) with respect to the proper system, from the transformation of the components of the $\mathbf{g}_{(a_0)}$, one will have:

$$(g^4_{(a_0)})_0 = \rho\beta c \frac{U_0}{m_0 c^2} \frac{\sqrt{1-\beta^2}}{\sqrt{1-\beta^2}+1} \cdot \frac{\beta'}{c\sqrt{1-\beta'^2}},$$

$$(g^1_{(a_0)})_0 = 0, \quad (g^2_{(a_0)})_0 = 0,$$

$$(g^3_{(a_0)})_0 = -\rho\beta c \frac{U_0}{m_0 c^2} \frac{\sqrt{1-\beta^2}}{\sqrt{1-\beta^2}+1} \cdot \frac{1}{\sqrt{1-\beta'^2}}.$$

For $\beta' = 0$, one will naturally recover:

$$(g^4_{(a_0)})_0 = 0.$$

For $\beta' = \beta$, one will recover the values of the system of the observer:

$$(g^4_{(a_0)})_0 = -\rho\beta \frac{U_0}{m_0 c^2} \frac{\beta}{\sqrt{1-\beta^2}+1} = \rho \frac{U_0}{m_0 c^2} (\sqrt{1-\beta^2} - 1)$$

and

$$(g^3_{(a_0)})_0 = -\rho\beta c \frac{U_0}{m_0 c^2} \frac{1}{\sqrt{1-\beta'^2}+1}.$$

11. CONCLUSION. NUMERICAL VALUES

In summary, the set of calculations that we just developed seems to us to prove that the passage to geometrical optics in Dirac's theory leads us to Weysenhoff's point mechanics, in which the existence of spin appears, rather than Pauli's point mechanics without spin. It then seems that we cannot exclude *a priori* the possibility of exhibiting the proper magnetic (or electric) moment of a Dirac particle.

Meanwhile it is difficult to find any convincing reasons to choose between the Pauli viewpoint and the one that we have developed above. In reality, the difficulty in the problem comes from the uncertainties in the application of the WKB method that were pointed out in paragraph 2 of this chapter. Indeed, if one writes:

$$\psi_k = a_k e^{\frac{2\pi i}{h} S} = a_k e^{\frac{2\pi i}{h} \int \frac{m_0 c^2}{\eta} dt + \dots} e^{\frac{2\pi i}{h} \int U dt},$$

and if one assumes that the proper moments are proportional to the $h / 2\pi i$ then the quantity $\frac{2\pi i}{h} \int U dt$ will have order zero in $h / 2\pi i$, and one will be led to write, with Pauli:

$$\psi_k = b_k e^{\frac{2\pi i}{h} \int \frac{m_0 c^2}{\eta} dt + \dots},$$

with:

$$b_k = a_k e^{\frac{2\pi i}{h} \int U dt} = b_k^{(0)} + \frac{h}{2\pi i} b_k^{(1)} + \dots + \left(\frac{h}{2\pi i} \right)^n b_k^{(n)} + \dots$$

On the contrary, if one assumes that even though the proper moments have order $h / 2\pi i$, the term U is very large, so that $\frac{2\pi i}{h} \int U dt$ will not be negligible compared to $\frac{2\pi i}{h} \int \frac{m_0 c^2}{\eta} dt$, then one will be led to apply the WKB method, as we did in paragraph 5 of this chapter. Now, that way of looking at things amounts to saying that one considers the quotient $K = \frac{\eta U}{m_0 c^2}$ to be non-negligible compared to unity, although its square can be considered to be negligible. Naturally, supposing that K is negligible compared to unity amounts to neglecting the existence of proper moments, and that would bring us back to the Pauli viewpoint. However, if we suppose that K is small compared to unity, but not negligible compared to it, then we will recover Weysenhoff's point mechanics, as we have seen, and Pauli's argument would not appear to be convincing.

By contrast, Bohr's argument, which is based upon the uncertainty relations, indeed shows the impossibility of resolving the decision, but only in the case of velocities that are small compared to the velocity of light. That impossibility does not seem to exist, in principle, for particles that are animated with velocities close to c ($h \gg 1$).

We now demand to know what numerical values can be attributed to the quantities ε , m_0 , and η in order for one to be able to observe a phenomenon that takes the form of a deviation that is caused by the action of the gradient of a magnetic field on the magnetic moment, which would conform to the picture of point mechanics.

Three conditions must be realized: First, in order for one to be dealing with point

mechanics, the quantity $K = \frac{\eta U}{m_0 c^2}$ must be small, which will give us the first condition:

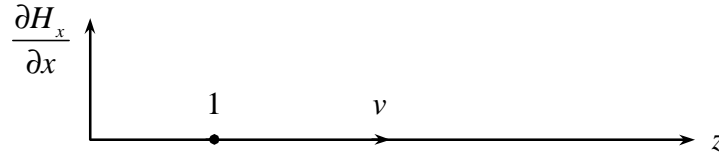
$$(I) \quad \frac{\varepsilon h \eta}{4\pi m_0^2 c^2} H \ll 1.$$

Next, in order for us to be able to have a group of waves that is point-like at our scale, it is necessary that the wave length must be very small at our scale (for example, equal to 10^{-4} cm or 10^{-5} cm), which will give us the second condition:

$$(II) \quad \frac{h}{m_0 c \eta} \sim 10^{-4} \text{ cm to } 10^{-5} \text{ cm.}$$

It will also be necessary that h / m_0c is not too small, as we showed in paragraph 2 of Chapter IX. We put c in the denominator, instead of v , in relation (III), because we supposed that $v \sim c$ in order to not revert to the conclusion of Bohr's argument.

Finally, it is further necessary that the deviation that is produced by the action of the gradient of the magnetic field on the proper moment must be observable. Suppose that the particle follows a path of length l along the Oz axis while being subject to a transverse field gradient $\partial H_x / \partial x$.



One will have:

$$\frac{d\pi_x}{dt} = \frac{\epsilon h}{4\pi m_0 c} \frac{\partial H_z}{\partial x},$$

and at the end of the trip:

$$\pi_x = \frac{\epsilon h}{4\pi m_0 c} \cdot \frac{1}{c} \cdot \frac{\partial H_z}{\partial x},$$

so if the deviation is assumed to be small then:

$$\tan \alpha \sim \alpha = \frac{\epsilon h}{4\pi m_0 c} \frac{\partial H_z}{\partial x} \frac{1}{\pi_z^{(0)}} = \frac{\epsilon h}{4\pi m_0 c} \frac{\partial H_z}{\partial x} \frac{1}{m_0 c \eta},$$

since $\tan \alpha = \pi_x^{(0)} / \pi_z^{(0)}$, or furthermore:

$$\alpha = \frac{\epsilon h H}{4\pi m_0^2 c^2 \eta} \frac{\partial H_z / \partial x}{H} l = \frac{K}{\eta^2} \cdot \frac{H'}{H} \cdot l.$$

Since $\eta > 1$, the first factor is very small, from (I), and in order to get an observable deviation, one must have:

$$\frac{\partial H_z}{\partial x} \cdot l \gg H_x,$$

which is a condition that seems difficult to realize.

While it is not impossible to obtain an observable deviation, in principle, it nonetheless seems difficult, in practice. Let us look at the numerical results.

We cannot take η to be less than 10^2 without reverting to Bohr's argument. With $\eta = 10^2$, we must take at least $m_0 = 10^{-35}$ g, or else we would have a wave length that would be too large ($\lambda > 10^{-6}$ cm) ⁽¹⁾. With $\eta = 10^2$, $m_0 = 10^{-35}$ g, we can take $\epsilon = 10^{-16}$ esu in

⁽¹⁾ The energy $m_0 c^2 \eta$ would then have order 10^{-12} ergs or 1 eV.

order to have $K \sim 10^{-2}$ with $H \sim 10^2$ Gauss. The deviation of $\alpha = \frac{K}{\eta^2} \frac{H'}{H} l = 10^{-6} \frac{H'}{H} l$ should not decrease below 10^{-2} (i.e., one cm per meter) if we are to get an observable phenomenon. Hence, $H' l \geq 10^4 H$. Upon taking the path length $l = 1$ m, which is already quite large, we will see that the value of the field H must vary by more than 1/10 mm. All of that is, without a doubt, quite difficult to realize.

If we assume the rather large value 10^2 for the ratio H'/H then the ratio:

$$\frac{\text{force due to the gradient}}{\text{Laplace force}} = \frac{h}{4\pi m_0 c} \frac{H'}{H} = \frac{\eta \lambda}{4\pi} \frac{H'}{H} \sim 10^{-3} \frac{H'}{H}$$

will have the value 1/10, which seems acceptable.

In brief: The givens $\eta = 10^2$, $m_0 = 10^{-35}$ g ⁽¹⁾, $\varepsilon = 10^{-16}$ esu $\sim 10^6 e$ will give the rigorously acceptable values $\lambda = 10^{-4}$ cm, $K = 10^{-2}$, with an even weaker deviation $\alpha = 10^{-2}$. We have reached the limit of possibility, and these values nevertheless seem to be the best ones that we can get.

In order to be able to exhibit the proper magnetic moment of a Dirac particle by a deviation of that kind, it will first be necessary that the particle should have precisely the characteristics that were indicated above, and even in that case, its detection would not be easy.

Jean Thibaud ⁽¹⁾ gave $\eta = 10^7$, $m_0 = 10^{-38}$ g, $\varepsilon = 10^{-10}$ to 10^{-14} esu for the electrinos that he thinks he has observed in his experiments, which will give the acceptable value of $\lambda = 10^{-6}$ cm for the wave length, and a ratio of forces of order 1 if H varies in value over 1 cm (i.e., if $H'/H \sim 1$). Unfortunately, one will then find that K has order 10^{11} to 10^{15} . The ratio K/η^2 will have order 1, and the deviation will be large, which is favorable. However, a value of K that elevated seems entirely irreconcilable with the validity of point mechanics, which Thibaud used in order to interpret his experimental results.

For values of η that are large enough to make $v \sim c$, the difficulty that presents itself here in regard to exhibiting the proper magnetic moment is a difficulty of a practical nature that relates to the effective possibility and precision of measurements, but that is not perhaps a theoretical impossibility *a priori*, as the arguments of Bohr and Pauli might indicate.

⁽¹⁾ In Thibaud's experiment, $H \sim 3000$ Gauss, $H'/H \sim 1/3$ in CGS.