THE MAGNETIC ELECTRON

(DIRAC'S THEORY)

BY

LOUIS DE BROGLIE

MEMBER OF THE INSTITUTE
PROFESSOR AT THE SORBONNE
NOBEL LAUREATE

Translated by
D. H. Delphenich

PREFACE

The theory of the electron that is due to Dirac is of great interest from several viewpoints. It constitutes the most perfect form that we currently possess for the wave dynamics of the electron. It reconciles relativistic and quantum ideas (at least, to some extent). It makes precise Uhlenbeck and Goudsmit's very fruitful hypothesis of an electrified corpuscle that is magnetic and rotating in a form that agrees with the principles of the new physics. Finally, it permits one to account for some important experimental facts that are concerned with the fine structure of spectra in the context of the anomalous Zeeman effect, and with that, in turn, it receives a beautiful confirmation. It seems useful to us to publish a general survey on that subject by editing a course that we taught in the most recent years at l'Institut Henri Poincaré.

In order to indeed show that Dirac's theory is not simply a game for theoreticians to play, but a timely way for us to explain some important facts, we believe that in the first part of the book we must survey some phenomena that have received a satisfactory interpretation from that theory, while at the same time showing that those phenomena defy any complete explanation in the old quantum theory or even wave mechanics in its initial form.

We have also reserved a special place in the first part of the book for a review of the general principles of the new mechanics and the conception of physical laws that arises from it. Indeed, it is impossible to comprehend Dirac's theory very well if one does not have those principles in mind.

In the discussion of the theory that defines the object of the second part, we have preserved the somewhat-asymmetric form of the equations that Dirac himself employed in the beginning without seeking to adopt notations that are more symmetric from a relativistic standpoint, as a number of authors have done since then. Indeed, that search for symmetry in form seems a little vain to us, since, as we have sought to show in the last chapter, Dirac's theory, despite the invariance of form of its equations under a Lorentz transformation, must make time play a special role in order to remain in agreement with the current general principles of quantum mechanics. At the end of the second part, we have devoted a chapter to a systematic overview of the entire theory that might aid the reader in understanding the harmony in it.

The third part of this book consists of the interpretation of some of the facts that were recalled in the first part by means of Dirac's theory. That is followed by a study of certain consequences of the fundamental equations that are a bit strange – in particular, its prediction of negative-energy states. We have presented these difficulties without advocating any solution to them. No matter what way of solving them that we are led to in the future, they deserve to be studied because they have their roots in the very basis of the theory itself.

We hope that this book will permit the reader to know not just the beauty of Dirac's theory, but also its utility in interpreting some experimental facts, as well as its gaps and weak points.

I would like to acknowledge Jean-Louis Destouches for having assisted me in the correction of the proofs.

<div align="right">LOUIS DE BROGLIE</div>

PART I

—

SUCCESSES AND FAILURES OF QUANTUM THEORY

AND THE ORIGINAL WAVE MECHANICS

TABLE OF CONTENTS

Page

CHAPTER V

Summary of the principles of wave mechanics

CHAPTER VI

Summary of the principles of wave mechanics (cont.)

CHAPTER VII

Relativistic form of wave mechanics for one wave function

CHAPTER VIII

Successes and failures of the wave mechanics of one wave function

PART TWO

Dirac's theory of the spinning, magnetic electron

General principles

CHAPTER IX

Pauli's attempt

CHAPTER X

Dirac's theory

CHAPTER XI

Relativistic invariance of the Dirac equations

CHAPTER XII

Density and current in Dirac's theory. Plane waves

CHAPTER XIII

The proper magnetism of the electron

CHAPTER XIV

The "electric and magnetic moment density" tensor

CHAPTER XV

**Matrices and first integrals in Dirac's theory.
Proper angular momentum of the electron**

CHAPTER XV

Systematic summary of the results obtained

PART THREE

**Applications of Dirac's theory.
Criticisms and various complementary notions.**

CHAPTER XVII

CHAPTER XVIII

Deduction of the Landé formula

CHAPTER XIX

Proper and orbital moment. Polarization of electronic waves

CHAPTER XX

Negative energy states in Dirac's theory

CHAPTER XXI

Schrödinger's zitterbewegung

Page

CHAPTER XXII

Some remarks on relativity and the new mechanics

_____

THE ATOMIC SPECTRUM OF HYDROGEN.
THE THEORIES OF BOHR AND SOMMERFELD.

___

1. The Balmer formula and the spectral terms of hydrogen. – The oldest-known series of lines in the visible light spectrum of hydrogen is the Balmer series. It consists of four principal lines (which are, in reality, small doublets, as we shall see later on). Here are the names and wave lengths of those four lines:

$$H_\alpha: 6.583 \text{ Å}, \quad H_\beta: 4.861 \text{ Å}, \quad H_\gamma: 4.340 \text{ Å}, \quad H_\delta: 4.102 \text{ Å}.$$

Balmer already arrived at a formula that gave the frequencies of those lines a half-century ago. That formula is:

$$v_m = R\left[\frac{1}{4} - \frac{1}{m^2}\right], \qquad m = 3, 4, 5, 6. \tag{1}$$

v_3 is the frequency of the line H_α, v_4 is that of H_β, etc. R is a constant that is called the *Rydberg constant*, and is approximately equal to 3.29201×10^{15}.

Some other line series in the non-visible hydrogen spectrum that were discovered later on obeyed analogous laws. Notably, one has the ultraviolet Lyman series, for which the frequencies of the lines are given by the formula:

$$v_m = R\left[1 - \frac{1}{m^2}\right], \qquad m = 2, 3, ... \tag{2}$$

and the infrared Paschen series, for which one has:

$$v_m = R\left[\frac{1}{9} - \frac{1}{m^2}\right], \qquad m = 4, 5, ... \tag{3}$$

One sees that all of those formulas have the general type:

$$v_m = R\left[\frac{1}{n^2} - \frac{1}{m^2}\right], \qquad n < m. \tag{4}$$

$n = 1$ for the lines of the Lyman series, $n = 2$ for those of the Balmer series, and $n = 3$ for those of the Paschen series.

One can elevate that statement to the status of a general law that is exhibited exactly for all spectral lines of all bodies. It is the *Ritz combination principle*, which one states in the form:

"The frequency of any spectral line is equal to the difference between two spectral terms that are characteristic of the emitting body,"

or, in another form:

"For any emitting body, one can construct a table of numbers that one calls *spectral terms,* such that the frequency of any spectral line of the body will be the difference between two of its spectral terms."

Formula (4) then shows that for hydrogen, the spectral terms have the form R / n^2, with $n = 1, 2, \ldots$, at least in absolute value.

A more precise experimental study of the lines of the Balmer series has shown, moreover, that those lines are, in reality, each composed of two very close lines. In other words, when one analyses the Balmer series with sufficient resolution, one will perceive that each of the lines that one originally considered to be simple is, in reality, a small doublet. The difference between the frequencies of the two components is the same for each of those doublets. Later on, we shall see how Sommerfeld could interpret that fine structure in the Balmer series.

2. Bohr's theory of the spectral terms of hydrogen. – In 1912, Bohr arrived at an interpretation for the spectral terms of hydrogen, and having done that, he founded the modern theory of the atom upon a new basis.

Somewhat before Bohr's theory, physicists (not without a certain amount of groping) followed the suggestion of Lord Rutherford and adopted a planetary model of the atom. According to that view, the atom of a simple body whose rank in the Mendele'ev series is *N* will be endowed with a central nucleus of positive charge *Ne*, where *e* is the elementary electronic charge $+ 4.77 \times 10^{-10}$ e. s. u. *N* electrons of charge – *e* orbit around that nucleus, in such a way that the atom will be collectively electrically neutral. Bohr had the idea of subjecting that model of the atom to calculation by applying the quantum laws that Planck introduced successfully in his study of black-body radiation. He assumed that a planetary electron in the atom can describe only certain motions around its central positive "Sun" that are predicted by classical mechanics. Those stable motions of the electrons correspond to "stationary states," during which no radiation will be emitted, contrary to the predictions of the classical electromagnetic theory. The emission of spectral lines can then take place only during the brief passage of the atom from one initial stationary state to another stationary state of less energy. What will the frequency of the line that is emitted during such a passage be? Bohr determined it by assuming that the energy that is lost by the atom is radiated in the form of just one quantum of light of value *hv* – or rather, in the form of just one *photon*, to employ the current language. Hence, if E_i and E_j denote the energies of the atom in the initial and final stationary states, resp., then the frequency v_{ij} that is emitted in the transition will be:

$$v_{ij} = \frac{E_i - E_j}{h} .$$

(5)

That formula explains the Ritz combination principle immediately, and we shall show that the spectral terms of an atom are equal to the energies of its various stationary states, divided by the Planck constant.

The essential problem is then to determine the energies of the stationary states. In order to do that, in his first paper, Bohr always assumed that the electron behaved like a point-like charge that obeyed Newton's laws of dynamics, but he restricted the number of possible motions by introducing Planck's quantization rules. In the era when Bohr was writing, one knew only how to quantize the periodic motions that were defined by just one variable q. The method of quantization for that case is the following one: If p is the Lagrange momentum that is conjugate to the variable q then one writes:

$$\oint p \, dq = n \, h \qquad (n, \text{ an integer}), \qquad (6)$$

in which the integral is taken over an entire cycle of the motion, and h is Planck's constant. Bohr was naturally led to suppose that the motions of the electron that satisfy the condition (6) are the stable motions that correspond to the stationary states of the atom.

The method permits one to easily calculate the energy of stable circular trajectories of a hydrogen atom, and in Rutherford's picture, that atom is composed of a nucleus of charge $+ e$ and a planetary electron of charge $- e$. If θ is the azimuth of the radius r then condition (6) will give:

$$mr^2 \, \dot{\theta} = n \cdot \frac{h}{2\pi} \qquad \left(\dot{\theta} = \frac{d\theta}{dt} \right). \qquad (7)$$

That expresses the idea that the angular momentum of the electron on a stable orbit is an integer multiple of $h / 2\pi$. On the other hand, since the laws of dynamics provide the relation:

$$mr^2 \, \dot{\theta} = \frac{e^2}{r^2}, \qquad (8)$$

one can easily find that the energy of the nth quantized circular motion is:

$$E_n = \tfrac{1}{2} mr^2 \, \dot{\theta} - \frac{e^2}{r} = - \tfrac{1}{2} mr^2 \, \dot{\theta} = - \frac{2\pi^2 me^4}{n^2 h^2}. \qquad (9)$$

Hence, if one sticks to the circular motions then the spectral terms of hydrogen must have the form:

$$\frac{E_n}{h} = - \frac{2\pi^2 me^4}{n^2 h^3}. \qquad (10)$$

From formula (5), the lines of hydrogen must have frequencies that are given by the general relations:

$$v_{nn'} = - \frac{2\pi^2 me^4}{h^3} \left(\frac{1}{n^2} - \frac{1}{n'^2} \right) \qquad (n' > n), \qquad (11)$$

and one will get back to formula (4), which is inferred from experiments, by setting:

$$R = \frac{2\pi^2 m e^4}{h^3}. \tag{12}$$

Now, the numerical calculation of the right-hand side of (12) shows that its value is indeed approximately equal to the experimental value of the Rydberg constant.

One can begin the same calculation again by supposing the one is dealing with an atom of atomic number N that has been ionized $N - 1$ times. One must then treat a problem that is identical with the problem of the hydrogen atom, except that the central charge will be Ne, instead of e ([1]). Upon repeating the calculation, one will easily find the following spectral terms, in place of (10):

$$\frac{E_n}{h} = - \frac{2\pi^2 m e^4}{n^2 h^3} N^2 = -\frac{RN^2}{n^2}. \tag{13}$$

Hence, the spectral terms are multiplied by the square of the atomic number. The simplest case is that of singly-ionized helium, for which one will have $N = 2$. The spectral terms and frequencies will then be quadrupled. Meanwhile, the experimental numbers indicate that the Rydberg constant does not have precisely the same value for H and He$^+$. Bohr could account for that difference by including the reaction of the electron on the nucleus.

3. Quantized energies of elliptic orbits. – Bohr's calculations that we recalled above cannot be regarded as being complete because the study of the motion of an electron around a nucleus is, in principle, a problem in two variables: viz., the radius vector and the azimuth. Eliminating the variations of the radius vector by considering only the circular trajectories is obviously only an artificial procedure. However, in order to treat the problem completely, one must first know how to write the quantum conditions for motions with several degrees of freedom. Here is how one arrives at that:

Let a system with n degrees of freedom, be defined by n variables q_1, …, q_n. If all variables admit the same period of variation T – i.e., if they take the same values at equal time intervals T – then the system will regularly take on the same configurations – i.e., it will be *periodic*. If each variable q_i has a period of variation T, and those periods are commensurable with each other, then the system will be *quasi-periodic*. For systems of the type that one had to quantize in the old quantum theory, it was always possible to choose the variables in such a fashion that they would form a system of n "separate variables"; i.e., each of the Lagrange momenta p_i could be expressed as a function of just the one corresponding coordinate q_i. Having chosen the variables in that way, H. A. Wilson and Sommerfeld showed that the quantization must be expressed by the n conditions:

$$\oint p_i \, dq_i = n_i \, h \qquad (n_i, \text{ an integer}) \qquad i = 1, 2, …, n \tag{14}$$

([1])   One often says that an atom of atomic number N that has been ionized $N - 1$ times is *hydrogenic*.

Sommerfeld appealed to that new statement in order to treat the problem of the hydrogen atom more completely by accounting for all of the elliptic motions. Let r be the radius vector, and let θ be the azimuth of the electron in its Keplerian trajectory. The kinetic energy T will be:

$$T = \tfrac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) , \tag{15}$$

and the Lagrange momenta will be, by definition:

$$p_r = \frac{\partial T}{\partial \dot{r}} = m \dot{r} , \qquad\qquad p_\theta = \frac{\partial T}{\partial \dot{\theta}} = m r \dot{\theta} . \tag{16}$$

The conditions (14) can be applied here, and one writes:

$$\int_0^{2\pi} m r^2 \dot{\theta} \, d\theta = n_1 \, h, \qquad\qquad \oint m \dot{r} \, dr = n_2 \, h. \tag{17}$$

Now, from the areal theorem, $m r^2 \theta$ is the constant angular momentum in a central field. The first condition (17) then gives:

$$m r^2 \dot{\theta} = n_1 \frac{h}{2\pi}, \tag{18}$$

and coincides with Bohr's condition (7) for circular orbits.

In order to calculate the second integral (17), we must write the expression for the energy:

$$E = \tfrac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) - \frac{e^2}{r} = \frac{1}{2m} \left[\frac{(m r^2 \dot{\theta})^2}{r^2} + m^2 \dot{r}^2 \right] - \frac{e^2}{r} . \tag{19}$$

Hence, upon taking (18) into account, one will infer that:

$$m \dot{r} = p_r = \pm \sqrt{2m \left(E + \frac{e^2}{r} \right) - \frac{n_1^2 \, h^2}{4\pi^2 r^2}} , \tag{20}$$

which is a formula that shows the separation of variables.

In the course of motion, the radius vector r oscillates between the values r_1 and r_2 that annul the radical (20), because p_r must be real. Upon supposing that $r_1 < r_2$, one can write:

$$\oint p_r \, dr = 2 \int_{r_1}^{r_2} \sqrt{2m \left(E + \frac{e^2}{r} \right) - \frac{n_1^2 \, h^2}{4\pi^2 r^2}} \, dr , \tag{21}$$

because one must take the + sign in (20) when r increases and the – sign when it decreases. Sommerfeld calculated the integral (21) by Cauchy's method of residues, and

found that it had the value $- \mid n_1 \mid h + \dfrac{2\pi^2 m e^4}{\sqrt{2m \mid E \mid}}$. Upon equating that to $n_2 \, h$, one will easily find that:

$$E_{n_1 n_2} = - \frac{2\pi^2 m e^4}{(\mid n_1 \mid + n_2)^2 h^2} \, . \tag{22}$$

That gives the quantized energy of the stationary state that corresponds to the quantum numbers n_1 and n_2 . Since $\mid n_1 \mid$ and n_2 are positive whole numbers or zero (which cannot both be zero, as one easily sees), one can set:

$$\mid n_1 \mid + n_2 = n \qquad (n = 1, \, 2, \, \ldots), \tag{23}$$

and the formula (22) will give the same energy levels as Bohr's original theory. In other words, the fact that one is considering elliptical orbits will not lead to any new spectral terms. The introduction of two degrees of freedom cannot explain the fine structure of the Balmer series by itself.

4. Sommerfeld's fine-structure theory. – In order to explain the fine structure of the hydrogen atom, Sommerfeld had the idea of attempting to employ relativistic mechanics, instead of classical mechanics. That would be justified by the remark that the speed of the electrons in the internal orbits of the Bohr atom must be comparable to the speed of light.

In relativistic dynamics, the kinetic energy of the electron has the expression:

$$T = m_0 \, c^2 \left[\frac{1}{\sqrt{1-\beta^2}} - 1 \right], \tag{24}$$

in which m_0 is the mass of the electron at rest, and β has the usual significance:

$$\beta = \frac{v}{c} = \frac{1}{c}\sqrt{\dot{r}^2 + r^2\dot{\theta}^2} \, . \tag{25}$$

However, here, one must no longer define the momentum p_i that is conjugate to the variable q_i to be equal to the derivative of T with respect to $\dot{q}_i$. One must introduce the relativistic Lagrangian:

$$L = - m_0 \, c^2 \, \sqrt{1-\beta^2} - U, \tag{26}$$

in which U is the potential energy, and pose the definition:

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{27}$$

In the problem of the hydrogen atom, the potential energy will not depend upon the q_i, and one will have:

$$p_r = \frac{\partial L}{\partial \dot{r}} = \frac{m_0 c^2}{2\sqrt{1-\beta^2}} \cdot \frac{\partial \beta^2}{\partial \dot{r}} = \frac{m_0 r}{\sqrt{1-\beta^2}} \,,$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \frac{m_0 c^2}{2\sqrt{1-\beta^2}} \cdot \frac{\partial \beta^2}{\partial \dot{\theta}} = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} \,.$$

(28)

Here, the quantum conditions (14) will then be:

$$\int_0^{2\pi} \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} \, d\theta = n_1 \, h, \qquad \oint \frac{m_0 \dot{r}}{\sqrt{1-\beta^2}} = n_2 \, h. \qquad (29)$$

The detailed study of the trajectory that we just carried out here showed that the electron describes an ellipse whose perihelion is rotating. In other words, at each instant, the trajectory is tangent to an osculating ellipse that turns slowly in its plane. The radius vector oscillates between two values r_1 and r_2, but the time that it takes to describe the cycle $r_1 \rightarrow r_2 \rightarrow r_1$ (viz., the period of the variable r) is a little bit longer than the time that the azimuth takes to increase by 2π (viz., the period of the variable θ). The orbit does not close precisely, so the motion will be quasi-periodic.
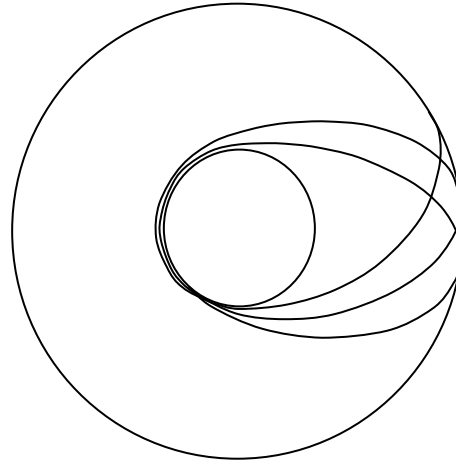
Figure 1.

The momentum p_θ is furthermore the angular momentum around the center, and one easily proves that it is once more constant here; i.e., that the areal theorem is always valid. The first condition (29) then gives:

$$p_\theta = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} = n_1 \frac{h}{2\pi} \,. \qquad (30)$$

The total energy is the sum of the internal energy $m_0 c^2$ of the kinetic energy and the potential energy. It is then equal to:

$$W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} - \frac{e^2}{r}. \tag{31}$$

Upon taking (28) and (25) into account, one will easily verify the formula:

$$W = c \sqrt{m_0 c^2 + p_r^2 + \frac{p_\theta^2}{r^2}} - \frac{e^2}{r}. \tag{32}$$

Let E denote the energy W minus the term $m_0 c^2$; E will then be the energy that one defines in classical mechanics. Replace W in (32) with $E + m_0 c^2$ and solve it for p_r. It will become:

$$p_r = \sqrt{A + 2\frac{B}{r} + \frac{C}{r^2}}, \tag{33}$$

with the notations:

$$A = \frac{E^2}{c^2} + 2m_0 E = m_0 c^2 \left[\left(1 + \frac{E}{m_0 c^2} \right)^2 - 1 \right],$$

$$B = \frac{E e^2}{c^2} + m_0 c^2 E, \tag{34}$$

$$C = \frac{e^4}{c^2} - p_\theta^2 = \frac{e^4}{c^2} - \frac{n_1^2 h^2}{4\pi^2}.$$

Sommerfeld then introduced what one calls the "fine structure constant" into the calculation:

$$\alpha = \frac{2\pi e^2}{hc} \qquad (\alpha^2 = 5.2 \times 10^{-5} \text{ c. g. s.}). \tag{35}$$

One can then write:

$$C = -\frac{n_1^2 h^2}{4\pi^2} \left[1 - \frac{\alpha^2}{n_1^2} \right]. \tag{36}$$

An application of the residue theorem permits one to establish that one has:

$$\oint \sqrt{A + \frac{2B}{r} + \frac{C}{r^2}} \, dr = -2\pi i \left(\sqrt{C} - \frac{B}{\sqrt{A}} \right). \tag{37}$$

Upon equating the right-hand side of (37) to $n_2 h$, conforming to (29), and replacing A, B, and C with their values, Sommerfeld found, after a simple calculation:

$$1 + \frac{E}{m_0 c} = \left[1 + \frac{\alpha^2}{[n_2 + \sqrt{n_1^2 - \alpha^2}]^2} \right]^{-1/2}, \tag{38}$$

which is a formula that rigorously gives the energy E of the stationary state that is defined by the quantum numbers n_1 and n_2 .

Since the quantity α^2 is very small, a first approximation will consist of neglecting the terms of degree higher than one in α^2. As one would expect, one then comes back to formula (22) from the non-relativistic theory, and one does not find any fine structure. A better approximation would consist of keeping the terms in α^4 and writing:

$$\left[1 + \frac{\alpha^2}{[n_2 + \sqrt{n_1^2 - \alpha^2}]^2} \right]^{-1/2} = 1 - \frac{\alpha^2}{2(|n_1| + n_2)^2} - \frac{1}{2} \frac{\alpha^4}{(|n_1| + n_2)^4} \left(\frac{1}{4} + \frac{n_2}{|n_1|} \right). \tag{39}$$

Upon substituting that value into (38), one will find that:

$$E = - \frac{2\pi^2 m_0 e^4}{(|n_1| + n_2)^2 h^2} \left[1 + \frac{\alpha^2}{(|n_1| + n_2)^2} \left(\frac{1}{4} + \frac{n_2}{|n_1|} \right) \right]. \tag{40}$$

The third term between brackets explains the existence of a fine structure in the Balmer series, because it depends upon $|n_1|$ and n_2 separately, and not just on $|n_1| + n_2$.

We shall now change the notations slightly. We call the number $n = |n_1| + n_2$ the "total quantum number," and the number $k = |n_1|$ the "azimuthal quantum number." It is obvious that one can characterize each quantum energy level by the numbers n and k, instead of the numbers n_1 and n_2 . Formula (40) can now be written (upon introducing the Rydberg constant):

$$E_{nk} = - \frac{Rh}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{k} - \frac{3}{1} \right) \right]. \tag{41}$$

In the old quantum theory, one assumed that the azimuthal quantum number k could never take the value 0; for the circular trajectories, one has $n_2 = 0$ or $n = k$, and for the elliptic trajectories $0 < k < n$. From formula (41), each stationary orbit is characterized by an energy E_{nk} that no longer depends upon just n, but also upon k. However, since α^2 is very small compared to unity, the various spectral terms that correspond to the same value of n will be very close to each other, and one will indeed get the fine structure of the lines that was predicted by Bohr's theory.

It is obvious that the Bohr spectral term that corresponds to a given value of n will decompose here into n terms that are close to each other, since one will have n values (1, 2, …, n) for k for a fixed n. It is appropriate to remark that the range of the closely-

spaced terms will become smaller as n increases, due to the presence of n_2 in the denominator of the term in α^2.

Consider the Balmer series. In the first approximation, the line frequencies are given by the formula:

$$v = R \left[\frac{1}{2^2} - \frac{1}{n^2} \right], \qquad\qquad n = 3, 4, \ldots \qquad (12)$$

In the second approximation, from Sommerfeld, one must replace the spectral term $\dfrac{R}{2^2}$ with $\dfrac{R}{2^2}\left[1+\dfrac{\alpha^2}{2^2}\left(\dfrac{2}{k}-\dfrac{3}{4}\right)\right]$ and similarly, the spectral term $\dfrac{R}{n^2}$ with $\dfrac{R}{n^2}\left[1+\dfrac{\alpha^2}{n^2}\left(\dfrac{n}{k}-\dfrac{3}{4}\right)\right]$. There is then one fine structure with constant width in the Balmer series that is provided by the doubling ($k = 1, 2$) of the first fixed spectral term and another fine structure with a width that decreases when one goes up in the series that is due to the complexity in the second variable spectral term. That second fine structure is practically unobservable, because it is too fine. The first one corresponds to a decomposition of each of the lines that were predicted by the Balmer formula into a doublet with constant width for the entire series, and is equal to:

$$\Delta v_{\mathrm{H}} = \frac{R}{2^2}\left[1+\frac{\alpha}{2^2}\left(\frac{2}{1}-\frac{3}{4}\right)\right] - \frac{R}{2^2}\left[1+\frac{\alpha^2}{2^2}\left(\frac{2}{2}-\frac{3}{4}\right)\right] = \frac{R\alpha^2}{16}. \qquad (43)$$

The numerical calculation of $\dfrac{R\alpha^2}{16}$ gives 0.365 cm^{-1} upon replacing frequencies with wave numbers. That value is in very good agreement with the experimental number. One then obtains an interpretation for the existence of doublets in the Balmer series by completing Bohr's theory with the introduction of relativity.

If one recalls Sommerfeld's calculation by supposing that one is no longer dealing with the hydrogen atom, but with an atom of atomic number N that has been ionized $N - 1$ times then one will recover the energy of the stationary state that is characterized by the numbers n and k:

$$E_{nk} = -\frac{RN^2h}{n^2}\left[1+\frac{\alpha^2 N^2}{n^2}\left(\frac{n}{k}-\frac{3}{4}\right)\right]. \qquad (44)$$

The relativistic correction in the bracket will then be multiplied by N^2, and the width of the doublets, by N^4. For ionized helium ($N = 2$), the width of the doublets in the series that corresponds to the Balmer series must then be 16 times larger than in the Balmer series itself. One then imagines that the study of the doublets in the spectrum of He$^+$ can serve to verify Sommerfeld's theory; Paschen's verification of that fact was quite satisfactory.

Sommerfeld's theory of fine structure then gave very good results for H and He$^+$. It could then be successfully extended to an interpretation of an important class of doublets in the X-ray spectrum. However, one then perceives that for X-rays, and even for the

simple hydrogen spectrum, the application of formula (41) will meet up with a grave difficulty. We shall explain that in detail in Chapter III.

_____

CHAPTER II

SUMMARY OF NOTIONS ABOUT THE OPTICAL SPECTRA OF DOUBLETS AND THEIR INTERPRETATION

___

1. Rydberg formula and line series

For some time now, the success of the Balmer formula and some formulas of the same type for hydrogen has led spectroscopists to see whether it would be possible to find analogous formulas for elements other than hydrogen. Since the combination principle is valid for all optical lines, that would amount to finding an expression for the spectral terms of a general element that would generalize the expression R / n^2 that is valid for hydrogen.

Rydberg showed that, in the first approximation, the spectral terms of an element can be written in the form:

$$\frac{R}{(n+\Delta)^2},\qquad(1)$$

in which R is the same constant as for hydrogen. In the expression (1), the number n is a positive whole number, and Δ is a non-whole number that is capable of taking several different values for each element. The Rydberg formula does not represent the spectral terms very exactly: Ritz proposed a more exact expression in which the whole number n and the quantity Δ appeared in the same way. Without going into the more exact expression for the spectral terms as functions of n and Δ here, we shall assume that each spectral term can be expressed with the aid of those two quantities, and it can, in turn, be represented by the notation (n, Δ). For a given value of n, Δ is, in general, capable of taking on several different values; the same n will then correspond to several spectral terms. Spectroscopists have made it a habit of denoting the possible values of Δ by a sequence of letters: $s, p, d, f, g, h, \ldots$, and upon constructing a table of spectral terms, one will see that such a table has the following appearance to it:

$$
\begin{array}{llllll}
(1, s) & (2, s) & (3, s) & (4, s) & (5, s) & \ldots\ldots \\
       & (2, p) & (3, p) & (4, p) & (5, p) & \ldots\ldots \\
       &   & (3, d) & (4, d) & (5, d) & \ldots\ldots \\
       &   &   & (4, f) & (5, f) & \ldots\ldots \\
       &   &   &   & (5, g) & \ldots\ldots
\end{array}
\qquad(2)
$$

Therefore, for $n = 1$, Δ will have just the value s. For $n = 2$, Δ can have the values s or p. For $n = 3$, Δ can have the values s, p, d, etc. More generally, when n increases by a unit, the number of possible values to Δ will also increase by unit.

As for hydrogen, the frequencies of the lines of the same spectral series are always obtained by taking the difference between the constant term that is characteristic of the

series and a variable term. Here are the frequencies of the four series that have been recovered in all spectra and are particularly familiar to spectroscopists:

Principal series $\qquad v_n = (1, s) - (n, p) \quad n = 2, 3, \ldots \quad \ldots\ldots$

Diffuse series, or $\qquad v_n = (2, s) - (n, d) \quad n = 3, 4, \ldots \quad \ldots\ldots$
1st second. series

Narrow series, or $\qquad v_n = (2, p) - (n, s) \quad n = 3, 4, \ldots \quad \ldots\ldots \qquad (3)$
2nd second. series

Bergmann series, or $\quad v_n = (3, d) - (n, f) \quad n = 4, 5, \ldots \quad \ldots\ldots$
fundamental series

Since the spectral terms always diminish in value when n increases, one will see that as one goes up in a spectral series, one will find lines whose frequencies approach the constant spectral term that characterizes the series more and more closely. That spectral term can then be called the *limit of the series.*

One should note a peculiarity of the formulas for the series (2) that one recovers for all of the series that are obtained under the usual conditions: If one ranks the possible values for Δ in the order s, p, d, f, g, etc., then the values of Δ that enter into a formula of the series will always have values that are immediately close. If one denotes the values of Δ by 1, 2, 3, …, instead of $s, p, d, \ldots$, then one can say that: *When one passes from one term to the other in a formula for a series, Δ will vary by one unit, more or less.* That is what one calls a *selection rule*, which shows that, at least under the usual conditions of emission, a great number of combinations of the spectral terms will not correspond to any actual observable lines.

2. The doublet spectra of alkali metals. – The table of spectral terms (2) is only a rough first approximation; unfortunately, reality is hardly that simple. In particular, progress in spectroscopy has shown that the lines that were considered to be simple in the preceding Rydberg-Ritz model are, in reality, composed of a group of neighboring lines that constitute doublets or triplets, or multiplets, more generally. Since we cannot carry out a complete presentation of that truly complicated question, we shall confine ourselves to an examination of the spectra of the alkali elements, in which the lines form doublets.

The study of the alkali metal spectra has shown that for those elements, most of the spectral terms in the Rydberg-Ritz model are doubled. More precisely, the value s of the quantity Δ will always remain unique, but the values $p, d, f, g, \ldots$ will all be double: There are two very close values p_1 and p_2, two very close values d_1 and d_2, etc. There is then good reason to replace table (2) with the following table:

$$
\begin{array}{cccccc}
(1, s) & (2, s) & (3, s) & (4, s) & (5, s) & \ldots\ldots \\
 & (2, p_1) & (3, p_1) & (4, p_1) & (5, p_1) & \ldots\ldots \\
 & (2, p_2) & (3, p_2) & (4, p_2) & (5, p_2) & \ldots\ldots \\
 & & (3, d_1) & (4, d_1) & (5, d_1) & \ldots\ldots \\
 & & (3, d_2) & (4, d_2) & (5, d_2) & \ldots\ldots \qquad (4)
\end{array}
$$

$$(4, f_1) \quad (5, f_1) \quad \ldots\ldots$$
$$(4, f_1) \quad (5, f_2) \quad \ldots\ldots$$
$$(5, g_1) \quad \ldots\ldots$$
$$(5, g_2) \quad \ldots\ldots$$

The usual optical series are obtained by combining spectral terms in which Δ takes two neighboring values from the list s, p, d, f, g, … That is the selection rule that was pointed out before. However, we shall find another supplementary one here. For example, consider a diffuse series of formulas:

$$v = (2, p) - (n, d), \tag{5}$$

and envision the lines of that series for which $n = 3$. *A priori*, that group of lines can be composed of the following four lines:

$$(2, p_1) - (3, d_1); \qquad (2, p_1) - (3, d_2); \qquad (2, p_2) - (3, d_1); \qquad (2, p_2) - (3, d_2). \tag{6}$$

Now, experiments prove that the second of those lines is never manifested under the usual conditions. That is a manifestation of a new selection rule whose precise statement we shall give after we have introduced the notion of quantum numbers into definition of the spectral terms in Table (4).

3. Interpretation of the Rydberg-Ritz model by Bohr's theory. – Bohr's theory gave excellent results for hydrogen and ionized helium, so it was quite natural to see to extend it to more complex atoms, and naturally, the primary goal to be reached in that attempt must be the interpretation of Table (2) of the Rydberg-Ritz spectral series.

One will encounter some grave difficulties when one seeks to extend Bohr's theory to atoms that contain more than one planetary electron. The dynamical problem becomes complicated, and the application of the quantum rules becomes uncertain. Meanwhile, the general analogy between the spectra of all the elements and the intervention of the Rydberg constant R in all of them leads one to think that the planetary model that is so useful for the hydrogen atom must be utilized to at least a certain degree for all of the other elements. In order to do that, we begin by assuming a very crude hypothesis: In an atom of atomic number N, consider $N - 1$ of the planetary electrons to be orbiting in the vicinity of the nucleus, so they form an "electronic shell" around it, while the Nth planetary electron, which is called the "optical electron," will have its orbit outside the shell. The transitions of the optical electron from one stationary state to another will determine the optical spectrum of the element. Thanks to the hypothesis of the electronic shell, one can assume approximately that the action of the nucleus of charge $+ Ne$ and the shell of charge $- (N - 1) e$ on the optical electron is equivalent, by compensation, to the action of a central charge $+ e$. One is then reduced to the problem of the hydrogen atom with just one quantum number. In the second approximation, with Sommerfeld, one seeks to account for the fact that the charge of nucleus and that of the shell do not compensate for each other to give precisely one unit. The trajectory of the electron will no longer be closed, so one must introduce a second quantum number k, and one can then

account for the transition from terms that have the Balmer form to terms of the form (*n*, Δ). Sommerfeld calculated (quite roughly, to be sure) the terms (*n*, Δ) that one will obtain in that fashion and recovered the formulas of Rydberg and Ritz. The theoretical attempt by Sommerfeld, and some other, more complicated, ones that one finds in the old books on quantum theory ([1]), have led one to make the various values of Δ (*s*, *p*, *d*, …) correspond with Sommerfeld's azimuthal quantum number in the following fashion:

$$k = \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad \ldots\ldots$$
$$\Delta = \quad s \quad p \quad d \quad f \quad g \quad h \quad \ldots\ldots \tag{7}$$

That will permit us to write any spectral term (*n*, Δ) in the form (*n*, *k*): For example, the term (2, *p*) is written (2, 2).

Since the azimuthal number *k* is always less than or equal to the total number *n*, but cannot take the value zero, one can explain why for *n* = 1, the quantity Δ can have only the value *s*, for *n* = 2, it can have only the values *s* and *p*, etc. The peculiarities of Table (2) are then found to have been explained completely. Moreover, by an argument that is based upon Bohr's correspondence principle, the old quantum theory showed that the only transitions that must take place effectively are the ones for which $\delta k = \pm 1$. That is, in fact, the selection rule that we pointed out for Table (2), since, from (7), a variation of *k* by plus or minus one unit would correspond to a displacement of one unit in the sequence of possible values of Δ.

4. Doublet spectra and the quantum number *j*. – By introducing the two quantum numbers *n* and *k*, the theories of Bohr and Sommerfeld then arrived at an interpretation of the spectral terms (2) in the Rydberg-Ritz model. However, we saw that this table is insufficient because the spectral terms that appear in it are multiplets, in reality. It is natural to think that in order to characterize each of the multiple spectral terms that correspond to the same term in Table (2), one must introduce a third quantum number. That is what had been done since the time of old quantum theory by introducing a third number *j* in an entirely empirical manner, in addition to the numbers *n* and *k*. Sommerfeld then gave the name of *internal quantum number* to that new concept, which is a name that is no longer very well justified today.

Without going into how the introduction of the quantum number *j* has permitted one to classify the complex optical multiplets here, we shall confine ourselves to studying the alkali doublets from that viewpoint. As we have seen, each spectral term (*n*, Δ) in that case will be double, in general. From the ideas that were presented in the last paragraph, that amounts to saying that the same value of the number *k* will correspond to two neighboring values for Δ, instead of a single one. Those two neighboring values of Δ that correspond to definite values of *n* and *k* must be characterized by different values of the quantum number *j*. One will then be led to attribute the two values of *j*:

$$j = k - 1 \pm \tfrac{1}{2} = l \pm \tfrac{1}{2} \tag{8}$$

([1]) In particular, see Léon BRILLOUIN: *L'atome de Bohr*, Paris, Presses universitaires (1931), ch. XII.

to the two neighboring spectral terms, if one introduces the notation:

$$k - 1 = l, \tag{9}$$

whose usefulness we shall see later on. Furthermore, one assumes that the number j cannot take on negative values, in such a way that for $k = 1$, one will have only the value $j = \frac{1}{2}$. One will then find the explanation for the uniqueness of the s terms in the table (4). Formula (8) leads immediately to the following table for the correspondence between the quantum numbers k and j and the values of Δ:

k	1		2		3		4		5			......
l	0		1		2		3		4			......
												......
j		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{9}{2}$		
Δ	s		p_1	p_2	d_1	d_2	f_1	f_2	g_1	g_2		......

$$\tag{10}$$

One can therefore represent each spectral term in Table (4) by the symbol (n, k, j), and one will get the following table in place of (4):

$$(1, 1, \tfrac{1}{2}) \quad (2, 1, \tfrac{1}{2}) \quad (3, 1, \tfrac{1}{2}) \quad (4, 1, \tfrac{1}{2}) \quad (5, 1, \tfrac{1}{2}) \quad \text{.......}$$
$$(2, 2, \tfrac{1}{2}) \quad (3, 2, \tfrac{1}{2}) \quad (4, 2, \tfrac{1}{2}) \quad (5, 2, \tfrac{1}{2}) \quad \text{.......}$$
$$(2, 2, \tfrac{3}{2}) \quad (3, 2, \tfrac{3}{2}) \quad (4, 2, \tfrac{3}{2}) \quad (5, 2, \tfrac{3}{2}) \quad \text{.......}$$
$$(3, 3, \tfrac{3}{2}) \quad (4, 3, \tfrac{3}{2}) \quad (5, 3, \tfrac{3}{2}) \quad \text{.......}$$
$$(3, 3, \tfrac{5}{2}) \quad (4, 3, \tfrac{5}{2}) \quad (5, 3, \tfrac{5}{2}) \quad \text{.......} \tag{11}$$
$$(4, 4, \tfrac{5}{2}) \quad (5, 4, \tfrac{5}{2}) \quad \text{.......}$$
$$(4, 4, \tfrac{7}{2}) \quad (5, 4, \tfrac{7}{2}) \quad \text{.......}$$
$$(5, 4, \tfrac{7}{2}) \quad \text{.......}$$
$$(5, 5, \tfrac{9}{2}) \quad \text{.......}$$

We have remarked that for the spectral terms of the alkali metals, there exists another selection rule of the same type as the one that is valid for Table (2) (i.e., $\delta k = \pm 1$), in addition to that rule, since, for example, the second of the four lines (6) cannot be produced under the usual conditions. That new rule is expressed with the aid of the number j by the formula $\delta j = \{- 1, 0 + 1\}$. We then have, in total, the following two rules:

$$\delta k = \begin{cases} +1 \\ -1 \end{cases}, \qquad \delta j = \begin{cases} +1 \\ 0 \\ -1 \end{cases}. \tag{12}$$

We shall now verify those rules on the four lines $(2p - 3d)$ that were given in (6). With the present notations, they are written:

$$(2, 2, \tfrac{3}{2}) - (3, 3, \tfrac{5}{2}), \quad (2, 2, \tfrac{1}{2}) - (3, 3, \tfrac{5}{2}),$$

$$\tag{13}$$

$$(2, 2, \tfrac{3}{2}) - (3, 3, \tfrac{3}{2}), \quad (2, 2, \tfrac{1}{2}) - (3, 3, \tfrac{3}{2}).$$

One has $\delta k = + 1$ for all four of them, but one of them – namely, $(2, 2, \tfrac{1}{2}) - (3, 3, \tfrac{5}{2})$ – will not satisfy the second rule in (12). In fact, it does not exist experimentally, either.

In the old quantum theory, the meaning of the number j remained unknown, and the attempts to interpret it proved to be insufficient. The Uhlenbeck-Goudsmit hypothesis of the spinning magnetic electron has revealed its true significance, and we shall see that emerge quite naturally in Dirac's theory.

_____

CHAPTER THREE

X-RAY SPECTRA AND
THE THEORIES OF BOHR AND SOMMERFELD

_____

1. Moseley's law and Bohr's theory

For the most part, the X-ray spectra present the same general character as optical spectra. The principle of combination is applicable to them. One encounters series that correspond to the combination of a fixed spectral term with a variable spectral term. One recovers selection rules for them.

In the first approximation, one can represent the X-ray spectral terms by the expression:

$$\frac{RN^2}{n^2}, \qquad n = 1, 2, \ldots \tag{1}$$

in which R is the Rydberg constant, and N is the atomic number of the simple body that is the emitter. That is *Moseley's law*. The presence of the constant R here indicates the close kinship between Röntgen spectra and optical spectra.

In reality, the form (1) of the spectral terms is only a rough first approximation, and the true X-ray spectral terms will present the same degree of complexity as the optical spectral terms do for the alkali metals. The doublets that one encounters here initially seemed to be explained by Sommerfeld's fine-structure theory, but a more careful examination will show that his theory is insufficient, and that things are ultimately put into order only by the theory of the magnetic electron, and in the Dirac form, in particular.

In the old, simplistic ideas of Bohr's theory, the emission of X-rays was linked to a reorganization in the electron shell. As we saw in the last chapter, Bohr considered the emission of light lines to be due to transitions from the outermost electron of one stable orbit to another, while the set of innermost electrons formed a sort of shell by the interlacing of its orbits. The emission of X-rays would then correspond to the modification that the internal system would be subjected to upon passing from one stable state to another. Naturally, the exact calculation of those stable states would be quite complicated, even in the old quantum theory.

One will be led to an approximate calculation (which is quite crude, moreover) in the following way: One assumes that the electronic orbits of the shell are individually characterized by one or more quantum numbers, which, in a certain sense, amounts to neglecting the interaction of the electrons and considering them separately. One then says that the electrons that possess equal or almost equal energies form a "band" (*couche*). Moreover, a band can be comprised of orbits of different types that are characterized by different sets of quantum numbers, but the energies of those orbits must be very close.

The calculations that were made for hydrogen showed us that the orbits of the same band must be characterized by the same total quantum number n. Experiments have

revealed the existence of bands in atoms that one has become accustomed to denoting by the consecutive letters of the alphabet *K*, *L*, *M*, etc. One accounts for that by setting $n = 1$ for the *K* band, $n = 2$ for the *L* band, etc.

The orbits of the *K* band have the smallest energy. From a well-known principle, it would then seem that the normal state of all electrons in the electron must be found in the *K* band. The study of X-ray spectra and the periodicity of the chemical properties in Mendele'ev's table prove that this is certainly not true. We must assume a sort of "band saturation"; i.e., assume that there exists a maximum number of electrons in each band. A critical examination of the facts that experiments provide on that subject permits us to state the following rule:

The maximum number of electrons that can belong to the band that is defined by total quantum number n is equal to $2n^2$.

Later on, we shall see how one can refine the distribution of electrons amongst the energy levels that belong to the same band.

We shall recall how Bohr and Kossel represented the emission of X-rays. An external agent (e.g., material particle or radiation) that arrives at an atom in the normal state can strip off one of the electrons in the shell and project it outwards. The atom will then be in an abnormal state; it will have suffered a "deep ionization." Let W_0 be the minimum normal energy of the atom then, and let W_1 be its energy after deep ionization.

The difference $W_1 - W_0$ was provided by the ionizing agent: We see that it corresponds (up to the factor $1 / h$) to the limiting frequency of a spectral series of X-rays. The atom then presents a free place in its shell, so a reorganization of it can be produced spontaneously, since an electron can abandon the place that it occupied to begin with and occupy the free place. Naturally, that transformation can be accomplished only if it corresponds to a diminution of the total energy of the atom, and must be accompanied by radiation that belongs to the X-ray domain.

Briefly: The origin of X-rays in this conception of the atom will be the transition of an electron from a certain band to another one of less energy in which a previous ionization has created an empty place.

After the reorganization of its shell, the atom will be found in a state of ionization that is less deep and will possess an energy W_2 that is between W_0 and W_1.

Imagine that second ionized state and remember that all of the electrons are considered to be similar: That new state will be identical to the one that the external agent would create if, instead of expelling the first electron, it expelled the second one, which is the one that was displaced during the reorganization. Hence, the energy $W_2 - W_0$ will represent the work that corresponds to the less-deep ionization. Bohr's rule then leads one to attribute the frequency:

$$v = \frac{1}{h}(W_1 - W_2) = \frac{1}{h}[(W_1 - W_0) - (W_2 - W_0)] \tag{2}$$

to the X-radiation that is emitted during the reorganization.

We construct the list of energies W_1, … of the atom in its various states of deep ionization by subtracting the energy W_0 from the normal state: We will then obtain what

we can call the energy levels of the atom when we start with the normal state. Upon dividing by h, we will get the X-ray spectral terms:

$$\frac{W_1 - W_0}{h}, \qquad \frac{W_2 - W_0}{h}, \qquad \dots, \qquad \frac{W_n - W_0}{h}, \qquad \dots \tag{3}$$

The X-ray spectral terms are then equal to the ionization work divided by h. The frequencies of the X-ray lines will then have the form:

$$v = \left[\frac{W_n - W_0}{h} - \frac{W_m - W_0}{h} \right]. \tag{4}$$

The lines in the same series correspond to the same W_n ; i.e., to the same level that the electron reaches at the time of the reorganization. The second term in (4) differs for different lines of the same series and tends to zero for lines of higher order: It is the frequency of the line when the order in the series is raised to a limit that is equal to the spectral term $(W_n - W_0) / h$ that characterized the series.

Moseley's law teaches us that in the sequence of elements, each of the spectral terms (3) will vary in the first (and very crude) approximation like the square of an atomic number. In order to recover that law in Bohr's theory, we assume that an electron of the shell can be regarded as being subject to a central field that is equal to $(N - z)\, e / r^2$, in which the term ze / r^2 crudely represents the repulsive action of the electron that are closer to the nucleus. We get the following general form for the spectral terms:

$$R \frac{N'^2}{n^2}. \tag{5}$$

z is small for the deep bands of heavy atoms, and one can make N coincide with N'. One then recovers the form (1) for Moseley's law. It is quite obvious that the calculation is hardly rigorous: However, Bohr's theory indicates that it does prove Moseley's law.

2. Summary analysis of X-ray spectra. – The first fact that struck the spectroscopists in the Röntgen domain was the following one: X-ray lines are very neatly divided into disjoint groups by the scale of frequencies or wave lengths. One first considers each of those groups of lines to be composed of just one series, and one then distinguishes those series by consecutive letters of the alphabet, starting with the letter K: e.g., K series, L series, M series, etc. The series will collectively displace towards the high frequencies when one passes from one element to a heavier one, and roughly as the square of the atomic number (Moseley). A more careful examination will then show that, in reality, the groups of lines that are first called the L series, M series, etc., will, however, decompose into sub-groups that form true series whose lines intermingle on spectrograms. There will then exist several L levels, several M levels, etc., of slightly differing energies. The K level is the only one that is unique.

Experiment has indicated that there exist 3 *L* levels, 5 *M* levels, and 7 *N* levels. Of course, each of those levels corresponds to a spectral term. One distinguishes the levels from their associated spectral terms by Roman numerals, where the lowest numerals correspond to the deepest levels; i.e., to the highest ionization work.

By way of example, here is the table of X-ray lines of the *K*, *L*, *M* series for the heavy elements:

Terms	K series	L series			M series				
	K	L_I	L_{II}	L_{III}	M_I	M_{II}	M_{III}	M_{IV}	M_V
L_I	…	…	…	…	…	…	…	…	…
L_{II}	α_2	…	…	…	…	…	…	…	…
L_{III}	α_1	…	…	…	…	…	…	…	…
M_I	…	…	η	l	…	…	…	…	…
M_{II}	β_2	β_4	…	…	…	…	…	…	…
M_{III}	β_1	β_3	…	…	…	…	…	…	…
M_{IV}	…	…	β_1	α_2	…	…	…	…	…
M_V	…	…	…	α_1	…	…	…	…	…
N_I	…	…	γ_5	β_6	…	…	…	…	…
N_{II}	γ_2	γ_7	…	…	…	…	…	…	…
N_{III}	γ_1	γ_3	…	…	…	…	…	…	…
N_{IV}	…	…	γ_1	β_2'	…	ε	…	…	…
N_V	…	…	…	β_2	…	…	γ	…	α_2
N_{VI}	…	…	…	…	…	…	…	β	α_1
N_{VII}	…	…	…	…	…	…	…	…	…
O_I	…	…	γ_{11}	β_1	…	…	…	…	…
O_{II}	…	γ_4'	…	…	…	…	…	…	…
O_{III}	δ_1	γ_4	…	…	…	…	…	…	…
O_{IV}	…	…	γ_2	…	…	…	…	…	…
O_V	…	…	…	β_5	…	…	δ	…	…
P_I	…	…	γ_2'	β_5'	…	…	…	…	…
P_{II}	…	…	…	…	…	…	…	…	…
P_{III}	…	γ_8	…	…	…	…	…	…	…

It is easy to interpret this table. The name of each line is inscribed at the intersection of a row and a column. Its frequency is the difference between the spectral term that is located at the head of the column and the spectral term that is located in the first column of the row. The lines of the same series are then found to be placed in the same column.

The unique *K* series is composed of a sequence of doublets: $\alpha_1 - \alpha_2$, $\beta_1 - \beta_2$, $\gamma_1 - \gamma_2$, … whose interval will get narrower. All of those doublets will dilate when the atomic number increases, and their frequency bandwidth will increase with the fourth power of the atomic number; one calls that kind of doublet a "regular doublet" or a "Sommerfeld doublet." The *L_I* series has a structure that is analogous to that of the *K* series; it is also composed of regular doublets whose width will diminish when one goes up in the series.

On the contrary, the *L_{II}* and *L_{III}* series present a very different aspect from the preceding ones. The homologous lines of those two series (which appear in the same row

in the table) form doublets with constant width from one end of the series ($\eta - l$, $\beta_1 - \alpha_2$, $\gamma_5 - \beta_6$, $\gamma_1 - \beta_2'$, $\gamma_{11} - \beta_7$, $\gamma_2' - \beta_5'$) to the other. That constant width is equal to $v_{LII} - v_{LIII}$ and varies with N^2 in the series of elements. Those doublets are once more "regular doublets.

One can summarize the laws of X-ray lines in general fashion by saying that: The X-ray spectral terms vary in magnitude with N^2 (viz., Moseley's law). The difference between the two consecutive spectral terms with the same name (L, M, …), the first of which is an even index (for example, $v_{LII} - v_{LIII}$) varies with N^4 and gives rise to regular doublets.

We further complete this discussion with the following remark: The difference between two spectral terms with the same name, the first of which is an odd index (for example, $v_{LI} - v_{LII}$), will vary in the sequence of elements in such a fashion that $\sqrt{v_{LI}}$ $- \sqrt{v_{LII}}$ will remain constant, and will define what one calls an "irregular doublet."

3. Classification of X-ray spectral terms. – The Bohr-Kossel model leads one to assume that the emission of an X-ray spectral series is a consequence of a deep ionization of an atom. If that ionization makes the energy of the atom pass from its normal value W_0 to the value W_n then the characteristic spectral term of the series will be $(W_0 - W_n) / h$. Assume that each electronic orbit can quantized separately. One can then presume that quantization will introduce three quantum numbers n, j, k, as it does for optical spectra. The deep ionization to which the energy W_n corresponds can then be represented by the symbol (n, j, k), which is composed of the three quantum numbers that define the orbit that is traversed by the expelled electron before ionization. It then results from this that we must make each X-ray spectral term correspond to one of the symbols (n, j, k) that were encountered before in optical spectra. Now the study of the X-ray series shows that they have entirely the same structure as the alkali metal spectra, in such a way that there exists a one-to-one correspondence between the X-ray spectral terms and the alkali spectral terms [Table (II) in the preceding chapter]. Here is that correspondence:

$$(1, 1, \tfrac{1}{2}) = K$$
$$1s$$

$(2, 1, \tfrac{1}{2}) = L_I$	$(3, 1, \tfrac{1}{2}) = M_I$	$(4, 1, \tfrac{1}{2}) = N_I$	$(5, 1, \tfrac{1}{2}) = O_I$	$(6, 1, \tfrac{1}{2}) = P_I$
$2s$	$3s$	$4s$	$5s$	$6s$
$(2, 2, \tfrac{1}{2}) = L_{II}$	$(3, 2, \tfrac{1}{2}) = M_{II}$	$(4, 2, \tfrac{1}{2}) = N_{II}$	$(5, 2, \tfrac{1}{2}) = O_{II}$	$(6, 2, \tfrac{1}{2}) = P_{II}$
$2p_1$	$3p_1$	$4p_1$	$5p_1$	$6p_1$
$(2, 2, \tfrac{3}{2}) = L_{III}$	$(3, 2, \tfrac{3}{2}) = M_{III}$	$(4, 2, \tfrac{3}{2}) = N_{III}$	$(5, 2, \tfrac{3}{2}) = O_{III}$	$(6, 2, \tfrac{3}{2}) = P_{III}$ (6)
$2p_2$	$3p_2$	$4p_2$	$5p_2$	$6p_2$
	$(3, 3, \tfrac{3}{2}) = M_{IV}$	$(4, 3, \tfrac{3}{2}) = N_{IV}$	$(5, 3, \tfrac{3}{2}) = O_{IV}$	

$$3d_1$$
$$4d_1 \qquad 5d_1$$
$$(3,\ 3,\ \tfrac{5}{2}) = M_V \qquad (4,\ 3,\ \tfrac{5}{2}) = N_V \qquad (5,\ 3,\ \tfrac{5}{2}) = O_V$$
$$3d_2 \qquad\qquad 4d_2 \qquad\qquad 5d_2$$

$$(4,\ 3,\ \tfrac{5}{2}) = N_{VI}$$
$$4f_1$$

$$(4,\ 3,\ \tfrac{5}{2}) = N_{VII}$$
$$4f_2$$

One must note that the construction of the O and P bands is not achieved in this table, even for heavy atoms, for lack of electrons.

Guided by the optical analogy, we now see that the K series has the spectral formula $(1,\ s) - (n,\ p)$: It is then a primary principal series that is composed like that of the alkali doublets that get narrower. The L series is a "secondary principal series" whose formula is $(2,\ s) - (n,\ p)$ and whose structure is analogous. The analogues for the K and L_I series are then found to be interpreted.

From our present viewpoint, the L_{II} and L_{III} series decompose into two sets of lines. The first set is composed of the regular doublets with constant width and simple components $\eta - l$, $\gamma_5 - \beta_6$, $\gamma_{11} - \beta_7$, $\gamma_2' - \beta_5'$, and is the analogue of the narrow series $(2,\ p) - (n,\ s)$ of the alkali metals. The second set of lines forms a diffuse series with the formula $(2,\ p) - (n,\ d)$ and refers to the most intense lines of the group L. They form doublets with constant width, but they present a fine structure that is due to the complexity of the d terms. One can, moreover, verify the selection rules $\delta k = \pm 1$, $\delta j = \{0,\ \pm 1\}$ in the table of X-ray lines that was given above, just as one does for optical spectra. The classification of X-ray lines by analogy with the doublet spectra of alkali metals, which is a classification that extends to the M and N series, provides a perfectly clear and coherent schema, and seems to impose the distribution of quantum numbers amongst the levels that was given by Table (6).

4. Theoretical interpretation. Sommerfeld's fine-structure formula. – We shall now see how one can seek to account for the allure of the X-ray spectral terms in the old quantum theory. The oldest and most simplistic image consists of regarding the atom of atomic number N as containing N planetary electrons that revolve around coplanar, concentric, circular orbits (e.g., the K circle, the L circle, etc.) In order to roughly account for the mutual action of the electrons, one supposes that it can be translated into the simple effect of a screen that has an apparent diminishing of the nuclear charge as a consequence. Hence, a K electron will be subject to a force $(N - k)\ e^2 / r^2$, an L electron, to a force $(N - l)\ e^2 / r^2$, etc., in which k, l, … are called the "screening numbers." The calculations that are valid for the hydrogen atom can then be applied here with no difficulty, and one will find the spectral terms:

$$\frac{-Rh\,(N-k)^2}{1^2}, \qquad \frac{-Rh\,(N-l)^2}{2^2},\ \dots; \tag{7}$$

i.e., roughly Moseley's law. This initial approximate theory is obviously insufficient for many reasons, and in particular, because it predicts only one L level, one M level, etc.

In order to explain the multiplicity of levels by bands, Sommerfeld once more introduced relativistic dynamics here, while considering all of the orbits to be circular or elliptic. We have seen that, by using that process, he found the spectral terms:

$$\frac{E_{nk}}{h} = - \frac{RN^2}{n^2}\left[1 + \frac{\alpha^2 N^2}{n^2}\left(\frac{n}{k} - \frac{3}{4}\right)\right] \tag{8}$$

for an atom of rank N that had been ionized ($N - 1$) times, in which α is the fine-structure constant, and k is the azimuthal quantum number, such that $0 < k \le n$. If one assumes that one can roughly represent the mutual repulsion of the electrons by a screening number then the *X-ray* spectral term that is characterized by the quantum numbers n and k will be given by:

$$\frac{E_{nk}}{h} = - \frac{R(N - z_{nk})^2}{n^2}\left[1 + \alpha^2 \frac{(N - z_{nk})^2}{n^2}\left(\frac{n}{k} - \frac{3}{4}\right)\right], \tag{9}$$

in which z_{nk} is the screening number that relates to the trajectory that is characterized by n and k. Since k can take n distinct values for a given value of n, formula (9) will predict 1 K-level, 2 L-levels, 3 M-levels, 4 N-levels, etc., which is still insufficient, as we saw in the preceding paragraph. Sommerfeld's theory, which is more complete than the original theory, is once more too restrictive then, since it does not introduce the quantum number j.

Nevertheless, despite its obvious inadequacy, Sommerfeld's theory seemed to score a great victory by its quantitative explanation for regular doublets. For example, take a regular doublet of the L series. It is provided by combining an M, N, etc. term with the terms L_I and L_{II}, respectively.

The two lines of the doublet then have frequencies of the form $v_{LII} - v_i$ and $v_{LIII} - v_i$. Their frequency difference δv_L (or width of the doublet) is $v_{LII} - v_{III}$. In his theory, Sommerfeld passed over the L_I level as temporarily inexplicable, and attributed the quantum numbers $n = 2$, $k = 1$ and $n = 2$, $k = 2$ to the L_{II} and L_{III} levels, respectively. Formula (9) then gives ([1]):

$$\delta v_L = v_{LII} - v_{III} = R\,\alpha^2\,\frac{(N - z_L)^4}{2^4}\left(\frac{2}{1} - \frac{2}{2}\right) = \frac{\pi^4 m e^8}{2c^2 h^5}(N - z_L)^4. \tag{10}$$

The width of the L doublets must then vary like $(N - z_L)^4$, and that is quite reasonably the experimental law, because z_L must be small in comparison to N for atoms that are not too light. If one sets $N = 1$, $z_L = 0$ then one will revert to the width Δv_H of the doublets in the Balmer series [formula (43) of Chapter I]. One will then have:

$$\delta v_L = \delta v_H (N - z_L)^4. \tag{11}$$

([1]) Upon letting z_{21} and z_{22} coincide and setting both of them equal to z.

Formula (11) is verified numerically quite well in experiments, on the condition that one must set $z_L = 3.5$, which is a reasonable hypothesis. The doublets of the series M and N are also predicted quite well an application of formula (9). That exact prediction of the regular doublets initially seemed to be a great success of Sommerfeld's fine-structure theory, but a grave objection then appeared, which we shall now discuss.

5. Inadequacy of Sommerfeld's fine-structure theory. – By introducing only the two quantum numbers n and k, Sommerfeld's theory did not give enough levels for the X-ray spectrum. We must then introduce the third number j, and the optical analogy constrains us to distribute the quantum numbers amongst the levels in the manner that is indicated by Table (6). From Table (6), the L_{II} and L_{III} levels have, in effect, the symbols $(2, 2, \frac{1}{2})$ and $(2, 2, \frac{3}{2})$; they will then have the same number for k (viz., 2) and differ by their number j. However, that would ruin the interpretation of the regular L doublets by formula (9), as Sommerfeld had done, because that formula assumes that the k numbers for L_{II} and L_{III} differ by unit. One might then believe (which would also be quite surprising) that the success of Sommerfeld's theory was purely fortuitous. The development of the more recent theories, and in particular, the ones that constitute the focus of this book, has since then shown that this fortuitous character was only apparent. It is, in fact, the introduction of relativity that permits one to correctly interpret the regular doublets, but on the condition that one must, at the same time, introduce the magnetic character of the electron, and it is the absence of the latter element from Sommerfeld's theory that is the source of its weakness.

We just saw that a careful examination of the X-ray spectra will show the inadequacy of Sommerfeld's fine-structure theory. However, even in the simple case of hydrogen, a deeper examination of the doublets in the Balmer series has shown that those doublets are not interpreted correctly by that theory. Consider the H_α line of the Balmer series: It comes from the transition from a stationary state for which $n = 3$ to a stationary state for which $n = 2$. In reality, that line is multiple, and if we utilize a mode of representation of the levels that is well-known today then Sommerfeld's theory will give us the following model for the fine structure of that line:
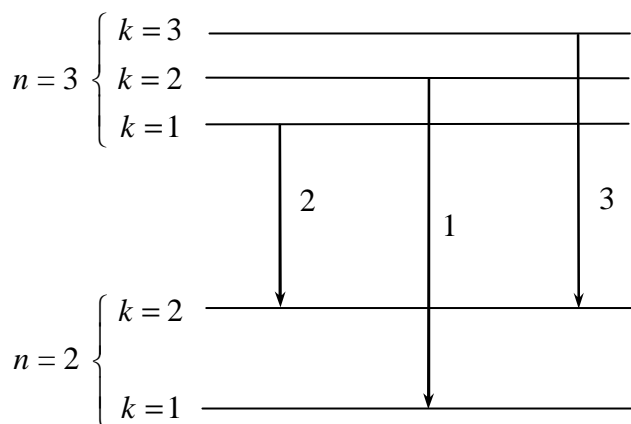
Figure 2.

Upon applying the selection rule $\delta k = 1$, one will predict the three lines that are indicated in the fine structure of H_α. The Sommerfeld doublet is composed of the line 1 and the set of two lines 2 and 3 with closely-spaced frequencies that are generally indiscernible.

Using the model with three quantum numbers n, k, j, one predicts the following fine-structure:

$$
n = 3 \begin{cases} k = 3 \begin{cases} j = 5/2 \\ j = 3/2 \end{cases} \\ k = 2 \begin{cases} j = 3/2 \\ j = 1/2 \end{cases} \\ k = 1 \quad\ j = 1/2 \end{cases}
$$

$$
n = 2 \begin{cases} k = 2 \begin{cases} j = 3/2 \\ j = 1/2 \end{cases} \\ k = 1 \quad\ j = 1/2 \end{cases}
$$
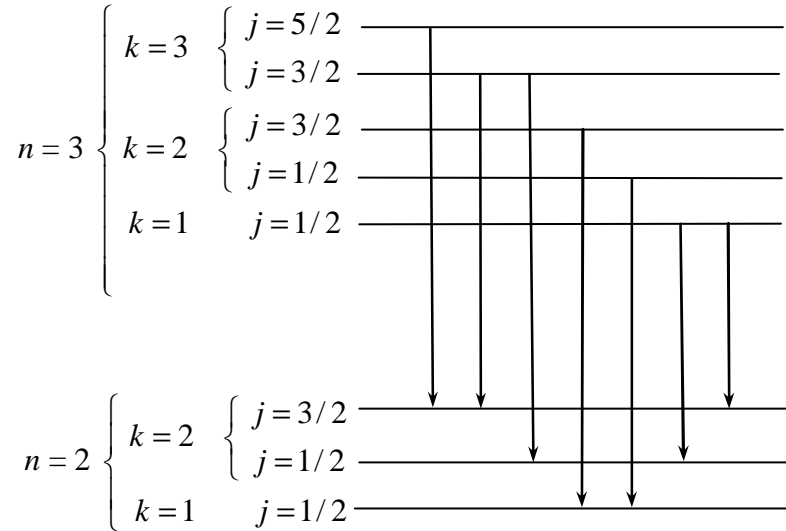
Figure 3.

The selection rules $\delta k = \pm 1$ and $\delta j = 0, \pm 1$ allow for the seven lines that are indicated in the figure above. However, for reasons that will become clearer later on, *for hydrogen,* one must consider levels that have the same j and different k as coinciding in the new theories of the magnetic electron. One will then have the following simplified model:

$$
n = 3 \begin{cases} k = 3, \quad j = 5/2 \\ k = 3, \quad j = 3/2 \\ k = 2, \quad j = 3/2 \\ k = 2, \quad j = 1/2 \\ k = 1, \quad j = 1/2 \end{cases}
$$

$$
n = 2 \begin{cases} k = 2, \quad j = 3/2 \\ k = 2, \quad j = 1/2 \\ k = 1, \quad j = 1/2 \end{cases}
$$
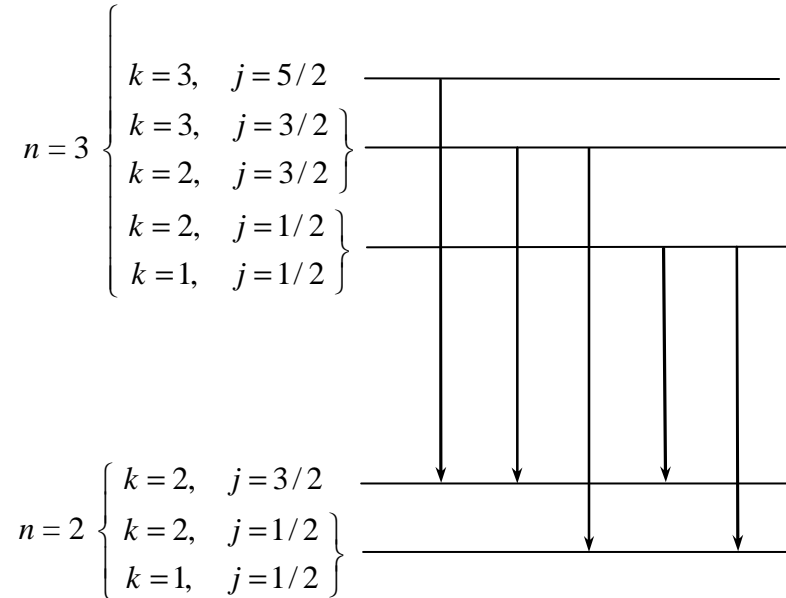
Figure 4.

We must then have five components to the fine structure of H_α, whereas Sommerfeld predicted only three. Now, the organized study of that fine structure has allowed us to assert that it involves more than three lines, which then confirms the new model, and not the Sommerfeld model. The study of the doublets of He$^+$ has allowed us to confirm that conclusion.

Therefore, even for hydrogen and ionized helium, Sommerfeld's theory fails, in the final analysis, at least in its original form.

6. Distribution of electrons amongst the levels. Stoner's rule. – As we have said, for a heavy atom, the various electrons distribute themselves into a certain number of bands K, L, etc. Those bands decompose into sub-bands or levels; hence, the L band will consist of the levels L_I, L_{II}, L_{III} .

We have seen that no band can possess more than a certain number of electrons; we spoke of a saturation of the bands. It is quite natural to think that there might exist a saturation of the bands. While being aided by the chemical properties of the elements and the character of their optical and Röntgen spectra, various authors (Bohr, Main Smith, Dauvillier and L. de Broglie, etc.) arrived at the study of the distribution of the electrons amongst the levels for the various elements. One will then have to know what the maximum number of electrons that each level can possess would be. On that subject, Stoner stated a rule that has been generally adopted by now: *The level that corresponds to the symbol* (n, k, j) *contains a maximum of* $2j + 1$ *electrons.*

Stoner's rule permits one to calculate the maximum number of electrons that belong to the band that is defined by a given value of the total quantum number n.

Indeed, since j can take the values $k - 1 \pm \frac{1}{2}$ for a given value of k, the levels:

$$(n, k, k - \tfrac{1}{2}) \quad \text{and} \quad (n, k, k - \tfrac{3}{2})$$

will possess a maximum of:

$$2\left(k - \tfrac{1}{2}\right) + 1 + 2\left(k - \tfrac{3}{2}\right) + 1 = 2\,(2k - 1)$$

electrons, resp., and the electrons of the band n will possess a maximum of:

$$2\sum_{k=1}^{n} (2k - 1) = 2\,[n\,(n+1) - n] = 2n^2 . \tag{12}$$

That is precisely the law that was pointed out in paragraph 1.

_____

CHAPTER IV

THE MAGNETIC ANOMALIES AND
THE SPINNING ELECTRON HYPOTHESIS

———

1. The gyromagnetic anomalies

Some simple electromagnetic considerations permit one to establish a general relation between the magnetic moment $\mathfrak{M}$ that is produced by the displacement of charge under the action of a central force and the constant angular momentum **M** that corresponds to the moment.

Consider a closed, planar trajectory that is traversed by a corpuscle of mass m_0 and electric charge ε under the action of a *central force* ([1]). From the areal theorem, the angular momentum $M = m_0 vr \sin \alpha$ is constant. If $d\mathcal{A}$ is the area that is swept out by the radius vector during the time dt then one will have:
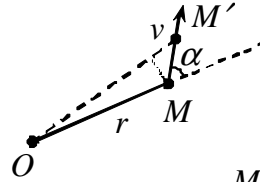
$$MM' = v\, dt$$
$$\text{area}: MOM' = \tfrac{1}{2} r \cdot v\, dt \sin \alpha$$

Figure 5.

$$d\mathcal{A} = \tfrac{1}{2}\, r \cdot v\, dt \sin \alpha = \frac{1}{2m_0} M\, dt. \qquad (1)$$

Since M is a constant, from the areal theorem, one will find upon integrating over a period T of the motion that:

$$\mathcal{A} = \frac{M}{2m_0} T, \qquad (2)$$

where $\mathcal{A}$ is the total area that is bounded by the trajectory.

On the other hand, the motion of the charge ε is equivalent to the existence of a current i. That current is, by definition, equal to the quantity of electricity that traverses a unit area normal to the trajectory every second. Since a charge ε passes through it during the T, one will have:

$$i = \frac{\varepsilon}{T}, \qquad (3)$$

———

([1]) We shall denote the mass of the corpuscle by m_0, instead of m, in order to avoid confusing that mass with the quantum number m later and in order to prepare us for the passage to relativistic equations.

upon assuming that ε is expressed in electromagnetic units. One infers from (2) and (3) that:

$$i = \frac{M\varepsilon}{2m_0\mathcal{A}}.$$ (4)

From the magnetic viewpoint, that current is equivalent to a leaf (*feuillet*) of strength i and area $\mathcal{A}$. The magnetic moment $\mathfrak{M}$ of that leaf will be:

$$\mathfrak{M} = i\mathcal{A} = \frac{\varepsilon M}{2m_0}.$$ (5)

That relation is valid in magnitude and direction for the vectors $\mathfrak{M}$ and $\mathbf{M}$. Hence:

$$\frac{\mathfrak{M}}{\mathbf{M}} = \frac{\varepsilon}{2m_0}.$$ (6)

One can prove in a more general fashion that if one considers an ensemble of corpuscles of the same mass m_0 and the same charge ε that form a system in a stationary state then the relation (6) will also be valid between the total magnetic moment $\mathfrak{M}$ of the motion of these charges and the constant total angular momentum $\mathbf{M}$ of the system ([1]).

From Bohr's theory, atoms are ensembles of electrons in stationary motion. One can thus apply relation (6) to atoms, with the condition that one must set $\varepsilon = -e / c$, where e is the unit positive charge that is commonly expressed in electrostatic units. One will thus have the formula ([2]):

$$\frac{\mathfrak{M}}{M} = -\frac{e}{2m_0c} \qquad (m_0 = \text{mass of the electron})$$ (7)

That fundamental formula leads to the idea of "Bohr magneton." Indeed, the old quanta theory always set the total moment of the quantity of motion for an atom equal to an integer multiple of $h / 2\pi$, so:

$$\mathfrak{M} = n\,\frac{h}{2\pi}\cdot\frac{e}{2m_0c} = n\,\frac{eh}{4\pi m_0c} \qquad (n = \text{positive or negative integer}).$$ (8)

The magnetic moment of the atom will thus always be an integer multiple of a certain unit that is called "the Bohr magneton," and is equal to:

([1]) The proof is due to Einstein. See the report of DE HAAS in *Atomes et Electrons* (*Rapports du Conseil de Physique Solvay de 1921*, Gauthier-Villars ed., Paris, 1923)

([2]) This formula is true *with* its sign, under the following conditions: One takes a system of *direct* axes and one defines the moments in such a way that the angular momentum of a corpuscle that spins around the z-axis in the direct sense (inverse to the hands of a clock) is directed in the positive sense of *oz*.

$$B = \frac{eh}{4\pi m_0 c}.\qquad\qquad(9)$$

The well-known Stern-Gerlach experiment exhibited the actual existence of the Bohr magneton.

Meanwhile, the relation (7) is not verified in a general and complete fashion. One must discover whether upon subjecting a magnetic bar to a magnetic field that bar will be put into rotation: Indeed, the theory suggests that the bar must take on a rotational motion such that ratio of its magnetic moment to its angular momentum is given by the relation (7). The phenomenon does, in fact, exist (the Einstein and de Haas experiment), but the ratio ($\mathfrak{M} / M$) was found to be equal to $e / m_0 c$, instead of $e / 2m_0 c$! Barnett found the same abnormal ratio by studying the inverse phenomenon (viz., the creation of a magnetic moment by the rotation of a bar). There was a grave difficulty in it that led to the idea of a proper magnetism of an electron, as we will confirm.

2. The normal Zeeman effect. – Another difficulty that suggests the magnetic electron hypothesis is the existence of anomalies in the Zeeman effect.

First, recall briefly the classical theory of the normal Zeeman effect that was once given by Lorentz. Consider the motion of a corpuscle of mass m_0 and charge ε in a uniform, magnetic field H. The corpuscle is subjected to a force that equal to:

$$F = \frac{\varepsilon}{c}[\mathbf{v} \cdot \mathbf{H}].\qquad\qquad(10).$$

Sir J. Larmor proved a very interesting and celebrated theorem on the motion that then results:

If one considers a reference system that rotates around the direction of a uniform, magnetic field with the constant angular velocity:

$$o = -\tfrac{1}{2}\frac{\varepsilon H}{m_0 c}\qquad (\varepsilon \text{ in e. s. u.})\qquad\qquad(11)$$

then the motion of the corpuscle in that reference system will be the one that one would have in a fixed reference system in the absence of a magnetic field when the other forces remain the same.

We apply that to an intra-atomic electron that is animated with a periodic motion of frequency v. If one creates a uniform, magnetic field then the same motion can be executed by the electron in the reference system that rotates with the Larmor precession velocity (11). The frequency $o / 2\pi$ that corresponds to that precession will be added or subtracted from the frequency of motion of the electron according to the relative orientation of the magnetic field and the orbit. It will then result that a material that emits radiation of frequency v in the absence of the magnetic field must also emit frequencies v

$+ \dfrac{1}{4\pi} \dfrac{\varepsilon H}{m_0 c}$ and $v - \dfrac{1}{4\pi} \dfrac{\varepsilon H}{m_0 c}$ in the presence of a uniform, magnetic field H. Upon examining the theory more closely, one will easily see that upon observing at a right angle to the magnetic field, one must see a line of frequency v that vibrates in the sense of the field and two lines that vibrate at a right angle with frequencies $v \pm \dfrac{1}{4\pi} \dfrac{\varepsilon H}{m_0 c}$, whereas when one observes them in the same sense as the field, one must see only the last two lines with circular vibrations in the opposite sense. This constitutes the normal Zeeman effect, which has indeed been observed in a certain number of cases, and which was discovered some thirty years ago, and which constitutes a brilliant verification of Lorentz's conception of the electron.

The old theory of quanta did not introduce anything that was essentially new in regard to the prediction of the Zeeman effect. Let an electron in an atom be in a stable state in the absence of an external, magnetic field; let W_0 denote the energy of that electron, and let $\mathfrak{M}$ denote the magnetic moment of its orbit, and assume that we have the right to consider one of the electrons of the atom thus-isolated. In the presence of a uniform, external, magnetic field H, the energy of the electronic orbit will be:

$$W_H = W_0 - (\mathfrak{M} \cdot \mathbf{H}), \tag{12}$$

or, by virtue of (6):

$$W_H = W_0 + \frac{e}{2m_0 c} (\mathbf{M} \cdot \mathbf{H}). \tag{13}$$

In order to define $(\mathfrak{M} \cdot \mathbf{H})$, we must consider the component of $\mathbf{M}$ along the field and that component must be a multiple of $h / 2\pi$, from the old quantum theory [1]. Formula (13) will therefore become:

$$W_H = W_0 + m \frac{ehH}{4\pi \, m_0 c}. \tag{14}$$

m is the quantum number that one calls the "magnetic quantum number."

Now consider the line that is due to the passage of the electron from the stable state of energy W_0 to a stable state of lower energy W_0'. In the absence of an external field, that line will have the frequency:

$$v_0 = \frac{W_0 - W_0'}{h}. \tag{15}$$

In the presence of a uniform field H, by virtue of (14), it will become:

$$v_H = \frac{W_H - W_H'}{h} = v_0 + (m - m') \frac{eH}{4\pi \, m_0 c}, \tag{16}$$

---

[1]   See Léon BRILLOUIN: *L'Atome de Bohr*, *loc. cit.*, pp. 167.

where $(m - m')$ can take on all positive or negative integer values, including zero. In order to recover the normal Zeeman effect, the old quantum theory, guided by the correspondence principle, assumed the selection rule $\Delta m = 0, \pm 1$. It then arrived at the same conclusions as the classical theory, and that was because the constant h was eliminated in the combination of formulas (14) and (15).

3. Anomalous Zeeman effect. Landé factor. – The theory of the normal Zeeman effect that was given in the preceding paragraph is verified only in a small number of cases. Most of the Zeeman effects are abnormal. We shall content ourselves by describing the abnormal Zeeman effect for the alkali elements, because they are the simplest ones, and they are also the only ones for which Dirac's theory, which has been powerless to account for the interaction of electrons, up to now, can be made complete.

The Zeeman effect for alkali metal is subject to the following general rules:

a) Homologous lines of the various alkali elements exhibit the same Zeeman effect.

b) Lines of the same spectral series exhibit the same decomposition (Preston's rule).

c) Lines that are displaced by the action of a magnetic field always define a figure that is symmetric in frequencies and polarization to the normal effect with respect to the original line.

d) The spectral interval (difference of frequencies) between a displaced line and the original line is always equal to the product of the normal Lorentz interval with a simple fraction (Runge's rule).

Therefore, for a certain line, the displaced components of the Zeeman effect will always be in the scale of frequencies with distances from the original line that are given by the expression $\pm \dfrac{s}{r}\dfrac{eH}{4\pi m_0 c}$, where r is a characteristic integer of the line in question (Runge denominator), and where s will take on a certain number of integer values, and is a number that determines the multiplicity of the decomposed line.

Runge's rule (*d*) can be interpreted in the following fashion: From the combination principle, the decomposition of a line translates into the decomposition of the spectral terms. One is then led to think that each spectral term must have its Runge denominator. Let r_1 and r_2 be the denominator of the two spectral terms. The Runge rule can be stated by saying that under the effect of the magnetic field, those terms will be modified by quantities $q_1 / r_1 \Delta v_{\mathrm{II}}$ and $q_2 / r_2 \Delta v_{\mathrm{II}}$, respectively; q_1 and q_2 are integer numbers, and Δv_{II} denotes the normal Lorentz interval $\dfrac{eH}{4\pi m_0 c}$. The line that is due to the combination of the terms considered is, in fact, subject to the spectral displacement:

$$\left(\frac{q_1}{r_1} - \frac{q_2}{r_2}\right)\Delta v_{\mathrm{II}} = \frac{q_1 r_2 - q_2 r_1}{r r_{21}}\Delta v_{\mathrm{II}}, \tag{17}$$

and can be written in the form $s / r \Delta v_{\mathrm{II}}$; we then get back to the statement *d*) of Runge's rule.

The manner by which the spectral terms decompose in the presence of a magnetic field was specified by Landé, who contributed much to the unraveling of abnormal Zeeman effect. He stated the following rules:

1. Each spectral term whose symbol is (n, l, j) (upon employing the notation $l = k + 1$) decomposes in a *weak* magnetic field into $2j + 1$ terms that are characterized by the *half-integer* magnetic quantum numbers: $- j, - (j - 1), \ldots, + (j + 1), + j$.

2. Upon taking the Lorentz normal interval $\dfrac{eH}{4\pi \, m_0 c}$ to be the unit, the frequency gaps in the decomposed terms with respect to the original term will be given by the following table:

$$
\begin{array}{llcccccc}
s & \text{terms}: \left[n,0,\tfrac{1}{2}\right] & & & -1 & +1 & & \\
p_1 & \text{terms}: \left[n,1,\tfrac{1}{2}\right] & & & -\tfrac{1}{3} & +\tfrac{1}{5} & & \\
p_2 & \text{terms}: \left[n,1,\tfrac{3}{2}\right] & & -\tfrac{6}{5} & -\tfrac{2}{3} & +\tfrac{2}{5} & +\tfrac{6}{5} & \\
d_1 & \text{terms}: \left[n,2,\tfrac{3}{2}\right] & & -\tfrac{6}{5} & -\tfrac{2}{5} & +\tfrac{2}{5} & +\tfrac{6}{5} & \\
d_2 & \text{terms}: \left[n,2,\tfrac{5}{2}\right] & -\tfrac{15}{5} & -\tfrac{9}{5} & -\tfrac{3}{5} & +\tfrac{3}{5} & +\tfrac{9}{5} & +\tfrac{15}{5}
\end{array}
\tag{18}
$$

This table shows that the Runge denominators are equal to 1 for the s terms, 3 for the p terms, 5 for the d terms, etc. In a more general fashion, the Runge denominator for the term (n, l, j) is $2j + 1$.

3. In the transitions that give rise to observable lines, the magnetic quantum number can vary only between $- 1, 0, +1$. For $\Delta m = \pm 1$, the emitted line will vibrate rectilinearly in the normal direction to the field when it is observed at a right angle to the field; for $\Delta m = 0$, it will vibrate parallel to the field when it is observed under the same conditions.

If the magnetic number m has the values that are given by rule 1 then Landé wrote down the gaps that are listed in table (18) in the form mg, where g is a number – viz., the *Landé factor* – that can possess the following values:

$$
\begin{array}{c|c|ccccccccc}
 & l & j = & \tfrac{1}{2} & \cdots\tfrac{3}{2} & \cdots & \cdots\tfrac{5}{2} & \cdots & \cdots\tfrac{7}{2} & \cdots \\
\hline
s & 0 & \cdots & 2 & & & & & & \\
p & 1 & \cdots & & \tfrac{2}{3} & \cdots\tfrac{4}{3} & & & & \\
d & 2 & \cdots & \cdots & \cdots\tfrac{4}{5} & \cdots & \cdots\tfrac{6}{5} & & & \\
f & 3 & \cdots & \cdots & \cdots & \cdots & \cdots\tfrac{6}{7} & \cdots & \cdots\tfrac{8}{7} &
\end{array}
\tag{19}
$$

One can summarize the table (19) by the formula:

$$
g = \frac{j+\tfrac{1}{2}}{l+\tfrac{1}{2}} = \frac{2j+1}{2l+1}. \tag{20}
$$

This is the value of the Landé factor for the term (n, l, j) of *an alkali metal*.

The spectral term of the alkali atom that has the value (n, l, j) in the absence of a field can take on the $2j + 1$ values:

$$(n, l, j)_H = (n, l, j)_0 + mg \frac{eH}{4\pi m_0 c} \qquad (21)$$

in the presence of a uniform field H, where g has the value (20), and m can take any half-integer value that is found between $-j$ and $+j$ ([1]).

The old quantum theory is just as incapable of explaining the intervention of the factor g as the classical theory was. We confirm that wave mechanics is not more fortunate here. It is only the introduction of the magnetism of the electron that explains the origin of the factor g, and Dirac's theory will allow us to predict the correct form with no objections.

Formula (21) is valid only for a weak magnetic field. But, what can be the reason for that? We say that a magnetic field that produces an anomalous Zeeman decomposition in an alkali doublet is weak if the displacement of the terms in the spectral sequence that the presence of the field produces is small in comparison to the gaps in the components of the doublet in the absence of the field. From that definition, the same magnetic field can be regarded as weak or strong depending upon the situation. When the field is weak in the sense that was just defined, formula (21) will be applicable. When it cannot be regarded as weak, one will find a more complicated phenomenon that obeys a law that was stated by Voigt, although we shall not insist upon it here. However, if the field is very strong – i.e., if the Zeeman displacement that it produces is large compared to the normal gap in the doublet – then one will once more get a simple phenomenon, namely, the *Paschen-Bach effect:* One will then observe the *normal* Zeeman decomposition that is centered on the center of gravity of the two components of the original doublet.

4. Hypothesis of the spinning, magnetic electron. – If one considers the law of the anomalous Zeeman effect and compares it to the classical Lorentz theory then one will perceive that in order to recover that theory, one must pose the relation:

$$\left| \frac{\mathfrak{M}}{M} \right| = g \ \frac{e}{2m_0 c} , \qquad (22)$$

in place of relation (7).

That will then lead us to think that the there exist angular momenta and magnetic moments in matter that are not related by relation (7). The anomalies of the gyromagnetic effect that were pointed out at the end of paragraph 1 suggest the same conclusion. One can therefore not assume that all of the magnetism of the atom is provided by the circulations of the electrons, when they are imagined to be point-like charges. The idea can then be presented in the spirit of attributing a proper magnetic

---

([1]) Preston's rule (*b*) is explained by remarking that the Zeeman decomposition is always independent of the total number n; only that number will vary now when one passes from one line to the other in the same series.

moment and a proper angular momentum to the electron itself that are related to each other by a relation that is different from (7). It is that idea that Uhlenbeck and Goudsmit proposed in an ingenious manner before the very development of the new mechanics itself.

Uhlenbeck and Goudsmit assimilated an electron of the classical type to a sphere of electricity in rotation around one of its diameters that possessed a angular momentum $M = \frac{1}{2}\frac{h}{2\pi}$ and a magnetic moment that is equal to a Bohr magneton $\mathfrak{M} = \frac{eH}{4\pi \, m_0 c}$, in such a way that one will have the value $\frac{e}{m_0 c}$ that is revealed by the Einstein and de Haas experiments for $\left|\frac{\mathfrak{M}}{M}\right|$. Since the ratio $\left|\frac{\mathfrak{M}}{M}\right|$ for the electron is therefore twice the normal ratio (7), one can speak of the "double magnetism" of the electron. After the success of the Uhlenbeck-Goudsmit hypothesis, various attempts were made to obtain a classical model of an electron that spins, but these attempts have lost much of their interest today, since the development of the new mechanics has forbidden us to consider the electron to be a small body that is very localized in space.

The hypothesis of the spinning, magnetic electron, by its appearance, has permitted us to glimpse the solution of some difficulties that we shall enumerate. First, take the question of the regular X-ray doublets. Uhlenbeck and Goudsmit assumed that the magnetic axis of the electron is always normal to the plane of its trajectory. Since there remain two possible senses for the proper angular momentum vector (which is frequently called the "spin" vector), each trajectory with quantum numbers n and k will correspond to two possibilities. One will then understand the necessity of introducing a new number j that is capable of taking on two distinct values for the given n and k in order to achieve the quantum determination of the stable trajectory. Since the total angular momentum is the sum of the angular momentum of the electron over all of its orbit, which is equal to $k\,\frac{h}{2\pi}$, and the is spin $\pm\,\frac{1}{2}\frac{h}{2\pi}$, one can write:

$$M_{\text{total}} = \left(k \pm \frac{1}{2}\right)\frac{h}{2\pi}, \tag{23}$$

so one will be tempted to set $j = k \pm \frac{1}{2}$. However, we will soon verify that the new mechanics leads us to replace k with $l = k - 1$; one thus imagines that the true formula that relates j to k is the formula that was pointed out before:

$$j = k - 1 \pm \tfrac{1}{2} = l \pm\tfrac{1}{2}. \tag{24}$$

The number j thus appears to be express the total angular momentum in units of $\frac{h}{2\pi}$. With the classical ideas, the small magnet that is formed by the electron displaces in the

Coulomb or quasi-Coulomb field of the nucleus and the shell. Everything then happens as if the small magnet were subjected to the magnetic field:

$$\mathbf{H} = -\frac{1}{c}\ [\mathbf{v} \cdot \mathbf{h}], \qquad\qquad (25)$$

where **h** is the Coulomb field. Formula (24) gives the action of an electrostatic field on a magnetic pole in motion with the velocity **v**. Since the field **H** is perpendicular to the planar orbit that is described by the small electronic magnet in the Coulomb field, the potential energy of that small magnet in the field **H** will be:

$$U = \pm\,\mathfrak{M}\,\mathbf{H}\,, \qquad\qquad (26)$$

where $\mathfrak{M}$ is equal to a Bohr magneton, according to the Uhlenbeck-Goudsmit hypothesis, and one must take the + sign or the – sign according to the sense of the moment $\mathfrak{M}$ that is normal to the trajectory with respect to the sense of its description. As a result of the existence of the potential energy (25), each level (n, k) of the Sommerfeld theory can be decomposed into two levels (n, k, j). From the calculations that were made by Uhlenbeck and Goudsmit, which were later reprised and perfected by Thomas and Frenkel, upon always appealing to the old quantum theory, one will be permitted to recover a law in N^4 for the doublets that are provided by the spectral terms whose number j differs by one unit, and thus makes the difficulty that was encountered by the original Sommerfeld theory disappear (see, paragraph 5 of the last chapter). However, those calculations are prone to some objections and arrive at good results only by means of artificial hypotheses, such as, for example, the substitution of the quantum number l for the quantum number k, which is a substitution that is unjustified when one employs the old quantum theory. It is certain today that one cannot treat intra-atomic problems by the methods of the old quantum mechanics, and that one must take recourse to those of wave mechanics. We shall not therefore insist upon the calculations of the original theory of the magnetic electron [1].

    The anomalous Zeeman effect and the Landé formula have received the beginnings of an explanation by the hypothesis of the spinning, magnetic electron. We shall always limit ourselves to the case of the alkali metals. The external, optical electron of an alkali element possesses a total angular momentum (i.e., orbital moment + spin) that is equal to $j = l \pm \frac{1}{2}$ times the unit moment $h\ /\ 2\pi$, upon assuming the substitution of l for k, somewhat arbitrarily.

    What will the potential energy of that optical electron be in the presence of an external, magnetic field **H**? With the old quantum theory, assume that the component M_H of the total angular momentum in the sense of the field **H** is of the form $m\,h\ /\ 2\pi$, where m is the magnetic quantum number, to which we impose the rule that it must take one of the half-integer values between $-j$ and $+j$. The total magnetic moment will thus form an angle with the field **H** whose cosine is $m\ /\ j$. If the electron does not have proper double magnetism then its energy in the field **H** will be:

---

  [1] See Léon BRILLOUIN, *loc. cit.*, chapter XVI.

$$W_H = W_0 - \mathfrak{M}_H \cdot \mathbf{H} = W_0 + \frac{e}{2m_0c} \, M_H \, H = W_0 + m \, \frac{eh \, H}{4\pi \, m_0c}. \tag{27}$$

One comes back to formula (14) and the normal Zeeman effect. However, by virtue of the double magnetism of the electron, one will have (upon always writing l, in place of k):

$$\mathfrak{M}_H = -\frac{e}{2m_0c} \cdot l \frac{h}{2\pi} \cdot \frac{m}{j} \mp 2 \frac{e}{2m_0c} \cdot \frac{1}{2} \frac{h}{2\pi} \cdot \frac{m}{j} = -\frac{e}{2m_0c} \cdot \frac{h}{2\pi} \cdot \frac{l \pm 1}{j} \cdot m, \tag{28}$$

and as a result:

$$W_H = W_0 - \mathfrak{M}_H \cdot \mathbf{H} = W_0 + m \, \frac{l \pm 1}{j} \cdot \frac{eh \, H}{4\pi \, m_0c}, \tag{29}$$

which is a formula that is equivalent to the empirical relation (21), under the condition that one must set:

$$g = \frac{l \pm 1}{j} = \frac{j \pm \frac{1}{2}}{l + \frac{1}{2}} = \frac{2j \pm 1}{2l \pm 1}. \tag{30}$$

This formula, which was obtained under some fairly arbitrary hypotheses, presents an air of uncontestable kinship with Landé's empirical formula (20).

The hypothesis of the magnetic electron that was somehow developed in the context of the old quantum theory thus gave some interesting results. However, today the success of the new mechanics has shown that the questions that relate to the electron must be posed in a completely different manner. We must therefore now summarize the fundamental concepts of wave mechanics. At the same time, we shall have to examine how we have sought (without much success) to introduce relativistic ideas into the original form of that new mechanics. That introduction has not permitted us to resolve the difficulties that were pointed out in the preceding pages. It is Dirac's theory that can smooth out those difficulties by simultaneously introducing the principle of relativity and the proper magnetism of the electron in the context of wave mechanics.

––––––––––––

CHAPTER V

SUMMARY OF THE PRINCIPLES
OF WAVE MECHANICS

___

1. Viewpoint of the new mechanics

In the old mechanics, one considered corpuscles or material points to be small objects of negligible dimensions that had a well-defined position in space at each instant. If a corpuscle is in motion then the set of its successive positions will constitute its trajectory. The classical equations of Newtonian dynamics (or the somewhat-modified equations of Einsteinian dynamics) allow one to predict the entire course of motion when one knows the forces that the corpuscle is subject to and certain initial conditions. The corpuscle is found to be attached to a certain number of quantities, such as its coordinates, its energy, the components of its quantity of motion, and those of its angular momentum with respect to a point, etc. The old mechanics attributed a well-defined value to them at each instant, and its equations allowed one to rigorously calculate the sequence of those values in the course of time.

The viewpoint of the new mechanics is completely different. Indeed, for it, the quantities that are attached to the corpuscle do not generally have well-defined values, in such a way that it is not possible to speak rigorously of a position at each instant or a trajectory. One can only assign a certain number of possible values that are attached to the corpuscle at each instant, each of which is affected with a certain probability: That must say that if one performs a precise measurement of the quantity in question at the instant considered then that measurement will give one of the values that were predicted to be possible, and the probability that one of those possible values will be the result of the measurement can be calculated in advance.

Hence, whereas the objective of the old mechanics was to predict the evolution from a given initial state of the quantities that are attached to the corpuscle rigorously and uniquely, the more modest objective of the new mechanics is only to calculate the possible values of those quantities at each instant, along with their respective probabilities. From the mathematical viewpoint, the difference between the two mechanics translates into the following fact: Whereas the old theory started with differential equations that permitted one to express the coordinates of material points as functions of time, the new one starts with a partial differential equation that has the form of an equation of wave propagation. We shall learn the form of that fundamental equation by restricting ourselves to the case of just one corpuscle that is placed in a known external field, because the general case of a system of corpuscles in interaction, which was easy to treat with the original wave mechanics, will not interest us here, since Dirac's theory has not arrived at a translation of that problem for the magnetic electron in a satisfactory way, up to now.

2. Constructing the non-relativistic equation of propagation. – We shall write the equation of propagation of the new mechanics in its initial non-relativistic form. That equation is obtained, in a sense, automatically by starting with the expression for energy in the old Newtonian mechanics. Let a corpuscle of mass m displace in a field that is derived from a potential U (x, y, z, t). The classical expression for the energy of the corpuscle is:

$$E = \tfrac{1}{2} m v^2 + U \ (x, y, z, t). \tag{1}$$

The quantity of motion is, by definition, the vector:

$$\mathbf{p} = m \ \mathbf{v} \tag{2}$$

whose components are $p_x = m v_x$, etc.

There then exists the relation:

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U \ (x, y, z, t) \tag{3}$$

between the energy and the components of the quantity of motion.

The right-hand side of (3) can be represented by H $(x, y, z, t, p_x , p_y , p_z)$: It is the Hamiltonian function that expresses the energy at each instant t as a function of the coordinates of the corpuscle and the components of its quantity of motion (or Lagrange momentum).

Here is how one obtains the equation of propagation of wave mechanics for the corpuscle in question. One replaces p_x in the Hamiltonian function with:

$$-\frac{h}{2\pi i}\frac{\partial}{\partial x}, \qquad p_y \text{ with } -\frac{h}{2\pi i}\frac{\partial}{\partial y}, \qquad p_z \text{ with } -\frac{h}{2\pi i}\frac{\partial}{\partial z}.$$

One thus obtains an operator:

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

that is called the *Hamiltonian operator*. One will then obtain the wave equations of the new mechanics by writing:

$$H \ (\Psi) = \frac{h}{2\pi i}\frac{\partial}{\partial t} . \tag{4}$$

$\Psi(x, y, z, t)$ is the wave function of the corpuscle, and that function is essentially complex. Upon specifying the form of the operator H, one will easily find the following expression for (4):

$$\Delta\Psi - \frac{8\pi^2 m}{h^2} U(x, y, z, t) = \frac{4\pi i m}{h}\frac{\partial\Psi}{\partial t}, \tag{5}$$

in which Δ is the well-known Laplacian symbol. Since the wave equation is of first order in time, that will permit one to calculate the form of the wave function at any instant when one knows its form at the initial instant.

In the very important case where U does not depend upon time (viz., a constant external field), the wave equations will admit monochromatic solutions; i.e., ones that depend upon time by only a factor of the form $e^{\frac{2\pi i}{h}Et}$. Such a monochromatic wave satisfies the equation:

$$\Delta\Psi + [E - U(x, y, z)]\,\Psi = 0, \qquad (6)$$

or *Schrödinger equation*, which is a degenerate form of (5).

In the even more special case in which U is zero (viz., no external field), one can once more write equation (6) with $U = 0$ for monochromatic plane waves, and in that case, one will get the following monochromatic plane-wave solution:

$$\Psi = a\,e^{\frac{2\pi i}{h}[Et - \sqrt{2mE}(\alpha x + \beta y + \gamma z)]}, \qquad (7)$$

in which a is a constant amplitude, and α, β, γ are the direction cosines of the direction of propagation. The wave (7) has a frequency $v = E / h$ and a wave length $\lambda = \dfrac{h}{\sqrt{2mE}} = h / mv$.

In the beginning, it was the monochromatic plane wave that wave mechanics associated with the free, uniform, rectilinear motion of a corpuscle of mass m, energy E, and quantity of motion mv.

3. New conception of the quantities that are attached to a corpuscle. – We just saw that the transition from the old mechanics to the new is effected by replacing the components of the quantity of motion with the operators $-\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$, etc. That is only a particular application of a general idea from the new mechanics, which is an idea that consists of substituting operators for all the quantities of classical dynamics. It is easy to specify how one defines those operators. We already know what the operators are that correspond to p_x, p_y, p_z. On the other hand, the operator that corresponds to the energy will be the Hamiltonian operator H that was defined above. One makes a coordinate of the corpuscle – x, for example – correspond to the operator x, which signifies "multiplication by x." All of the other mechanical quantities are quantities that are derived from $x, y, z, t, p_x, p_y, p_z$. Hence, whenever a quantity is expressed as an entire, rational function of the coordinates and the momenta, one can construct the corresponding operator ([1]). Hence, for example, the z-component of the angular momentum of the corpuscle with respect to the origin of the coordinates will be replaced by the operator:

([1]) One can have ambiguities that come from the order of factors, but we shall not address them here, because we will not encounter them.

$$M_z = xp_y - yp_x = -\frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right), \tag{8}$$

upon taking the direct axes and agreeing that a motion that consists of a rotation in the positive sense in the *xy*-plane will correspond to a positive angular momentum.

The operators that one is led to construct in wave mechanics that correspond to measurable mechanics quantities are complex operators, in general, and they will all belong to one particular class – viz., the Hermitian operators.

Here is how one defines the class of Hermitian operators: We shall first introduce a convention that will persist in all of what follows in this book: A certain letter will represent an operator or a function, and the same letter, when endowed with an asterisk, will represent the complex-conjugate quantity. Having assumed that convention, let *A* be an operator of the type that we just considered. If $d\tau$ denotes the spatial volume element *dx dy dz* then the operator *A* will be *Hermitian*, by definition, if one has:

$$\int f^* A(g)\,d\tau = \int g\,A^*(f^*)\,d\tau, \tag{9}$$

in which the integrations extend over all space, and the functions *f* and *g* of the coordinates are finite, uniform, and continuous in all space, and tend to zero at infinity sufficiently rapidly that the surface integrals that are obtained by parts on the left-hand side of (9) are zero. The operators that are envisioned in wave mechanics are all Hermitian. It is easy to verify that for every operator, and in particular, one can verify it for the operator M_z in (8), for example.

Along with their Hermiticity, the operators of wave mechanics always have another common character: They are linear – i.e., one always has:

$$A\,(\varphi_1 + \varphi_2) = A\,(\varphi_1) + A\,(\varphi_2), \tag{10}$$

and as a result:

$$A\,(c\,\varphi) = c\,A\,(\varphi). \tag{11}$$

One must make an important distinction between the operators that the new mechanics attaches to a corpuscle. The one kind are concerned with the set of three coordinates *x*, *y*, *z*, and are called *complete* operators.

The others are concerned with only two coordinates, and for that reason, they are called *incomplete* operators. For example, the Hamiltonian operator *H* is complete, while the operators p_x or M_z are incomplete. Later on, we shall see the importance of that distinction.

Briefly, in wave mechanics, we make any dynamical quantity that is attached to a corpuscle correspond to a linear, Hermitian operator. However, it is quite obvious that if one makes a precise measurement of one of those mechanical quantities then the result of that measurement will be expressed as a real number. As we said in paragraph 1, the goal of the new mechanics is, first of all, to say what the real numbers are that a precise measurement would yield as the values of a mechanical quantity. We can then deduce a list of real numbers that represent all possible results of a precise measurement of a quantity that is attached to a corpuscle from the Hermitian operator that the new

mechanics associated with that quantity. Now, that will be possible because all of the Hermitian operators of wave mechanics possess a sequence of "proper values" that are real numbers. We shall now explain what that means.

   4. Proper values and proper functions of a Hermitian operator. – Let A be a Hermitian linear operator. Write the equation:

$$A\,(\varphi) = \alpha\,\varphi, \tag{12}$$

in which α is a constant, and φ is a function of the coordinates x, y, z.

   By definition, we say *proper values of the operator A* to mean the values of the constant α for which equation (12) admits at least one solution $\varphi\,(x, y, z)$, which one calls a *proper function of the operator A*, and which enjoys the following properties ([1]): It is everywhere finite, uniform, and continuous, and the integral of the square of its modulus over all space is well-defined and convergent. Naturally, if the operator A depends upon time then the same thing will be true for its proper values and proper functions.

   We assume the existence of proper values of the linear, Hermitian operators of wave mechanics, but we shall prove that those proper values are necessarily real. Indeed, the conjugate equation to (12) can be written:

$$A^{*}\,(\varphi^{*}) = \alpha^{*}\,\varphi^{*}, \tag{12*}$$

and since A is linear, one will have:

$$\int_{D}\varphi^{*}A(\varphi)\,d\tau - \int_{D}\varphi A(\varphi^{*})\,d\tau = (\alpha - \alpha^{*})\int_{D}\varphi\varphi^{*}d\tau\,, \tag{13}$$

in which the integration extends over the entire domain D of the variables that enter into φ; i.e., into A. Now, the left-hand side of (13) is zero, since A is Hermitian. Since the integral in the right-hand side is essentially positive, one must have $\alpha = \alpha^{*}$, so α will be real.

   The set of real proper values of a Hermitian operator is called the *spectrum* of that operator. That spectrum will be *discontinuous* if the proper values are isolated, and *continuous* if they form a continuous set. A spectrum can even be partially-continuous and partially-discontinuous. We shall first argue with discontinuous spectra.

   Let α_i denote an isolated proper value: There exists at least one proper function $\varphi_i\,(x, y, z)$ that corresponds to it. The set of proper functions defines an *orthogonal system*, in the sense that if φ_i and φ_j are two proper functions that correspond to two distinct proper values α_i and α_j then one will have:

$$\int_{D}\varphi_i^{*}\varphi_j\,d\tau = 0. \tag{14}$$

   ([1])   Here, we consider only linearly-independent functions to be distinct.

Indeed, if α_i is real then one will have:

$$A\,(\varphi_j) = \alpha_j\,\varphi_j\,, \qquad A^*\,(\varphi_i^*) = \alpha_j\,\varphi_j^*\,, \qquad (15)$$

and as a result:

$$\int_D \varphi_j A^*(\varphi_i^*)\,d\tau - \int_D \varphi_i^* A(\varphi_j)\,d\tau \;=\; (\alpha_i - \alpha_j)\int_D \varphi_i^* \varphi_j\,d\tau\,. \qquad (16)$$

Since the left-hand side is zero by reason of the Hermiticity of A, and $\alpha_i - \alpha_j$ is non-zero, by hypothesis, equation (14) will be proved.

The proof breaks down for two linearly-independent proper functions that correspond to the same proper value. When that case presents itself, one says that there is *degeneracy*, and that the proper value is *multiple*. Let α_i be a multiple proper value that corresponds to p linearly-independent proper functions φ_{i1}, φ_{i2}, …, φ_{ip} . Since the operator A is linear, any linear combination of the φ_{i1}, …, φ_{ip} will be a solution of $A\varphi = \alpha_i \varphi$. One can then replace the p linearly-independent proper functions φ_{i1}, …, φ_{ip} with p linearly-independent linear combinations of those functions, and one will easily see that one can choose those linear combinations in such a fashion that they are mutually-orthogonal. In other words, when a proper value is multiple, the system of linearly-independent proper functions will be determined only up to a linear transformation, and one can profit from that partial indeterminacy in order to get a system of independent proper functions that are orthogonal. Therefore, one can always assume that the set of proper functions of a Hermitian operator is orthogonal.

The proper functions of a Hermitian operator are determined only up to a complex constant factor (and even when there is no degeneracy). In order to fix the modulus of that complex factor, one is accustomed to *normalizing* the functions φ_i ; i.e., to setting:

$$\int_D \varphi_i^* \varphi_j\,d\tau = \int_D |\,\varphi_i\,|^2\,d\tau = 1, \qquad (17)$$

which is an equation that will be meaningful, since $|\,\varphi_i\,|^2$ is summable. The proper functions will still have an arbitrary factor of the form $e^{i\alpha}$ once they have been normalized.

Upon introducing the symbol δ_{ij} that is equal to one if $i = j$ and to zero if $i \neq j$, one can summarize the formulas (14) and (17) with the formula:

$$\int_D \varphi_i^* \varphi_j\,d\tau = \delta_{ij}\,. \qquad (18)$$

All of the preceding formulas apply to the discontinuous spectrum. If the operator A possesses a continuous spectrum then any proper value α of that spectrum will correspond to a proper function that we write as $\varphi\,(\alpha,\,x,\,y,\,z)$, because when α varies continuously in the continuous spectrum, it would be just as natural to write it as a variable as it would be to write it as an index. The proper functions $\varphi\,(\alpha,\,x,\,y,\,z)$ are orthogonal to the proper functions of the discontinuous spectrum, if there is one. However, in order to avoid certain difficulties regarding convergence, it will be more convenient in the study of continuous spectra to consider, instead of the proper functions

$\varphi\ (\alpha,\ x,\ y,\ z)$ themselves, the expressions $\dfrac{1}{\Delta\alpha}\displaystyle\int_{\alpha}^{\alpha+\Delta\alpha}\varphi(\alpha,x,y,z)\,d\alpha$, which are called *proper differentials* that correspond to intervals $\alpha \rightarrow \alpha + d\alpha$ that are chosen to be as small as one desires in the domain of continuous variation of α. The use of proper differentials leads one to replace formula (18) with the formula:

$$\frac{1}{\Delta\alpha}\int_{D} d\tau\left[\int_{\alpha}^{\alpha+\Delta\alpha}\varphi^{*}(\alpha,x,y,z)\,d\alpha\right]\left[\int_{\alpha'}^{\alpha'+\Delta\alpha}\varphi(\alpha,x,y,z)\,d\alpha\right] = \delta_{\alpha\alpha'}. \qquad (19)$$

Before concluding this paragraph, we must further point out a very important property of the proper functions of a linear, Hermitian operator *A*: They define a *complete system*. That says that, under very broad conditions, a function of the variables that relate to *A* (which vary in the domain *D*) can always be developed into a series of proper functions of that operator. For example, if $f\ (x,\ y,\ z)$ is a function of three variables x, y, z then it can be developed quite generally in proper functions of a complete, Hermitian operator *A* in the form of:

$$f(x, y, z) = \sum_{i} d_{i}\varphi_{i}(x, y, z) + \int d(\alpha)\,\varphi(\alpha, x, y, z)\,d\alpha, \qquad (20)$$

in which the sum Σ is extended over the discontinuous spectrum, and the integral, over the continuous spectrum.

If we exhibit the proper differentials that correspond to the various intervals $\Delta\alpha$ of the continuous spectrum then we can replace (20) with:

$$f(x, y, z) = \sum_{i} d_{i}\varphi_{i}(x, y, z) + \sum_{\Delta\alpha} d(\alpha)\left[\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha}\varphi(\alpha, x, y, z)\,d\alpha\right]\Delta\alpha. \qquad (21)$$

Upon utilizing formulas (18) and (19), one will easily find that:

$$d_{i} = \int \varphi_{i}^{*} f(x, y, z)\,d\tau\,;$$

$$d\ (\alpha) = \int_{D}\left[\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha}\varphi^{*}(\alpha, x, y, z)\,d\alpha\right] f(x, y, z)\,d\tau. \qquad (22)$$

The quantities d_{i} and $d\ (\alpha)$ are 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 the Fourier integral represent simple special cases of that type of development.

5.   General principles of wave mechanics. – In paragraph 1 of this chapter, we said that the objective of the new mechanics is to calculate the possible values of the quantities that are attached to the corpuscle and their respective probabilities. We then learned how to associate a wave function $\Psi\ (x,\ y,\ z,\ t)$ with the corpuscle that is a solution

to equation (5), which is a wave function that we always assumed to have been "normalized" by the condition ([1]):

$$\int \Psi \, \Psi^* d\tau = 1. \tag{23}$$

We then make each quantity that is attached to the corpuscle correspond to a linear, Hermitian operator that permits us to define a set of real numbers, namely, its proper values, and a complete system of normalized, orthogonal functions, namely, its proper functions.

We are then in a position to state the following two fundamental principles of the new mechanics:

First principle: The possible values at the instant t of a quantity that is attached to a corpuscle – i.e., the possible results of a precise measurement of that quantity that is performed at the instant t – are the proper values of the linear, Hermitian operator A that corresponds to that quantity at the instant t.

Second principle: If a corpuscle has a certain solution Ψ (x, y, z, t) of its wave equation for its wave function then the probability that a precise measurement of the quantity that corresponds to the complete operator A at the instant t will prove a certain proper value is equal to the square of the modulus of the coefficient of the corresponding proper function in the development of the wave function Ψ in normalized, orthogonal proper functions of the operator A.

More explicitly, if the function Ψ is developed in proper functions of A in the form [which is analogous to (21)]:

$$\Psi = \sum_i c_i \varphi_i + \sum_{\Delta\alpha} c(\alpha) \left[\frac{1}{\Delta\alpha} \int_\alpha^{\alpha+\Delta\alpha} \varphi(\alpha) \, d\alpha \right] \Delta\alpha \tag{24}$$

then $|c_i|^2$ will give the probability of the proper value α_i, and $|c(\alpha)|^2 \Delta\alpha$ will give the probability that α has a value that is found in the interval $\alpha \to \alpha + \Delta\alpha$. Since the function Ψ is normalized, the total probability of all the possible hypotheses will indeed be unity, as one easily verifies. Naturally, the probabilities that are provided by the second principle are, in general, functions of time t; i.e., the instant at which one measures them.

If the operator A admits multiple proper values then the statement of the second principle must be completed: Let α_i be a multiple proper value that corresponds to p linearly-independent normalized, orthogonal proper functions φ_{i1}, …, φ_{ip} . The probability that one will find the value α_i for the quantity in question when one measures it will be:

$$|c_{i1}|^2 + |c_{i2}|^2 + \ldots + |c_{ip}|^2 \, ;$$

([1]) Later on, we shall prove that if the condition (23) is satisfied at a certain instant then it will be satisfied at every instant.

i.e., the sum of the moduli of the coefficients of φ_{i1}, …, φ_{ip} in the development of Ψ in proper functions of A. That probability is independent of the manner by which one chooses the p proper functions φ_{i1}, …, φ_{ip}, as it should be.

When the operator A is incomplete, the statement of the second principle must be modified. Indeed, the proper functions of A will not contain all three variables x, y, z then, and the c_i and the $c(\alpha)$ in the development (24) will obviously be functions of variables that are contained in the φ_i and $\varphi(\alpha)$. The probability of a proper values α_i cannot be $|c_i|^2$ then, which will be a quantity that will further depend upon certain variables. In order to obtain the probabilities, it will then be necessary to *integrate* the expressions that were pointed out above over the entire domain of the variables that enter into A. For example, if A depends upon only y and z then the c_i will depend upon x, and the probability of the value α_i will not be $|c_i|^2$, but in fact, $\int_{-\infty}^{+\infty} |c_i|^2 \, dx$. One will verify that this modification is, indeed, in agreement with the idea that the total probability of all the possible hypotheses must be unity.

We shall indicate some examples of applications of the general principles. A very simple example is the application to the Hamiltonian, which is, as we know, a complete operator. For the Hamiltonian operator, equation (12) can be written:

$$H(\varphi) = E\,\varphi, \tag{25}$$

if one writes E in place of α. One will have proper values E_i and proper functions φ_i. Those proper values and proper functions will depend upon time when A does; i.e., when the system is not conservative. A precise measurement of the energy can yield only one of the values E_i that relate to the instant of time t when the measurement was performed as a result, and the probability that one will get the value $E_k(t)$ is equal to the square of the modulus of the coefficient of the function φ_k in the development of the wave function Ψ of the corpuscle in proper functions of the energy at the instant t. That is what we called "the spectral decomposition principle" in some other discussions.

We now seek to apply our principles to the quantity that is the x-coordinate of the corpuscle. Equation (12) takes the form:

$$x \cdot \varphi = \alpha\,\varphi. \tag{26}$$

That equation can be regarded as being satisfied for any real value of α by the function $\delta(x - \alpha)$, which is the *Dirac function*, and it enjoys the following properties:

1. It is an even function of the argument $(x - \alpha)$.
2. The integral $\int f(x)\,\delta(x-\alpha)\,dx$ is zero if the integration interval does not contain the value $x = \alpha$, and is equal to $f(\alpha)$ for any integration interval that does contain that value.

Equation (26) will then admit a continuous spectrum that includes all real values of α between $-\infty$ and $+\infty$.

From the first principle, a measurement of the coordinate can then give no particular value between $-\infty$ and $+\infty$, *a priori* (as it must). Moreover, as one easily sees, the

proper differentials $\dfrac{1}{\Delta\alpha}\displaystyle\int_{\alpha}^{\alpha+\Delta\alpha}\delta(x-\alpha)\,d\alpha$ of that continuous spectrum define a complete, normalized, orthogonal system. Since one has, from the definition of the δ-function, that:

$$\Psi\,(x,\,y,\,z,\,t) = \int_{-\infty}^{+\infty}\Psi(\alpha,x,y,z,t)\,\delta(x-\alpha)\,d\alpha\,,\tag{27}$$

from the second principle, the probability that a measurement of the x coordinate will give a value that is found in the interval $\alpha \to \alpha + \Delta\alpha$ will be:

$$\int_{-\infty}^{+\infty}dz\int_{-\infty}^{+\infty}dy\,|\,\Psi(\alpha,x,y,z,t)^2\,|\,d\alpha\,.\tag{28}$$

It will then easily result that the probability for a simultaneous measurement of the three coordinates to give values that are found in the intervals $\alpha \to \alpha + \Delta\alpha$, $\beta \to \beta + \Delta\beta$, $\gamma \to \gamma + \Delta\gamma$ is $|\,\Psi\,(\alpha,\,\beta,\,\gamma,\,t)\,|^2\,d\alpha\,d\beta\,d\gamma$. That amounts to the same thing as saying that the probability that a measurement will permit one to localize the corpuscle in the volume element $dx\,dy\,dz$ around the point whose coordinates are x, y, z is $|\,\Psi\,(x,\,y,\,z,\,t)\,|^2\,dx\,dy\,dz$. That is what we called the "interference principle" in other presentations.

6. Quantities that are or are not simultaneously measurable. – One can deduce a very important consequence of the general principles that were stated in the preceding paragraphs: Two mechanical quantities can be measured simultaneously and precisely only if the corresponding operators A and B commute; i.e., only if one has $AB = BA$.

Indeed, if φ_i and χ_i denote the proper functions of A and B, respectively, and α_i and β_i denote their respective proper values then in order for a simultaneous measurement of the two quantities in question to be made precisely, it is necessary that one must be able to simultaneously attribute a certain value α_i to the first quantity and a certain value β_i to the second one with certainty. From the second principle, it will then be necessary that one must be able to write the wave function Ψ of a corpuscle in the form:

$$\Psi = c_i\,\varphi_i = d_i\,\chi_i\,.\tag{29}$$

If A is an incomplete operator then c_i can depend upon variables that do not enter into φ_i, and if B is incomplete then d_i can likewise depend upon variables that do not enter into c_i. One infers from the preceding equation that:

$$AB\,(\Psi) = AB\,(d_i\,\chi_i) = A\,(d_i\,\beta_i\,\chi_i) = \beta_i\,A\,(c_i\,\varphi_i) = \beta_i\,\alpha_i\,\Psi\tag{30}$$

and

$$BA\,(\Psi) = BA\,(c_i\,\varphi_i) = B\,(c_i\,\alpha_i\,\varphi_i) = \alpha_i\,B\,(d_i\,\chi_i) = \alpha_i\,\beta_i\,\Psi\,.\tag{31}$$

One must then have:

$$AB\,(\Psi) = BA\,(\Psi)\tag{32}$$

for any value of Ψ that has the form (29), which implies that $AB = BA$.

The simplest and most important example of quantities that are not simultaneously measurable is that of a coordinate and the corresponding component of the quantity of motion. Indeed, one has:

$$x\, p_x - p_x\, x = \frac{h}{2\pi i}\left(\frac{\partial}{\partial x}\, x \cdot - x \cdot \frac{\partial}{\partial x}\right) = \frac{h}{2\pi i}\,. \tag{33}$$

In other words, the operator $xp_x - p_x x$ is equivalent to the multiplication by $h / 2\pi i$. Hence, x and p_x do not commute, and as a result, a coordinate and the corresponding component of the quantity of motion cannot be measured precisely and simultaneously. A coordinate and its corresponding Lagrange momentum can be known at a given instant t only with certain uncertainties Δx and Δp_x, which cannot both be zero. One can show that one will always have:

$$\Delta x\, \Delta p_x \geq h, \tag{34}$$

at least, in order of magnitude. The inequality (34) and the two analogous inequalities for y and z constitute the *Heisenberg uncertainty relations*, upon which we have insisted for some time in other books ([1]).

––––––––––

([1]) *Introduction à l'étude de la Mécanique ondulatoire*, Hermann, Paris, 1930.
La théorie de la quantification dans la nouvelle Mécanique, Hermann, Paris, 1932.
In those two books, one will find a more detailed discussion of the principles that were presented in summary form in the present chapter, as well as in the following one.

SUMMARY OF PRINCIPLES OF WAVE MECHANICS (CONT.)

_____

1. Some definitions that relate to algebraic matrices

One calls a table of numbers that involves either a finite or infinite number of rows and columns a *matrix.* If the table is finite-dimensional then we shall assume that it is square. More generally, we could assume that it is rectangular, but that would be a useless complication for us here. Each number that enters into the table (viz. each *element* of the matrix) can be located with the aid of two indices that define the row and column to which it belongs, respectively. Therefore, let a_{ik} denote the matrix element that is found to be written in the table at the intersection of the ith row and the kth column: The matrix will then be collectively represented by A or $|\,a_{ik}\,|$. The elements a_{ii} with equal indices are situated along the diagonal of the matrix and are called *diagonal elements.* We say that two matrices A and B are *equal* and write $A = B$ if their elements with the same indices are all equal ($a_{ik} = b_{ik}$).

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 transformations of the type:

$$x'_i = \sum_j a_{ij} x_j , \qquad (1)$$

which are formulas that one can condense by writing the vectorial relation:

$$\mathbf{X}' = \mathbf{A}\,\mathbf{X}, \qquad (2)$$

and upon agreeing that the vector $\mathbf{AX}$ has the quantity $(\mathbf{AX})_i = \sum_j a_{ij} x_j$ for the component with the index i.

Formula (1) leads one to define the sum and product of two matrices that have its same number of lines and columns by the following conventions:

1. The sum of two matrices A and B is the matrix $A + B$ whose element with the indices ik is $a_{ik} + b_{ik}$.

2. The product of the matrix with the matrix A is the matrix AB whose ik element is $(AB)_{ik} = \sum_l a_{il} \cdot b_{lk}$.

It results from definition 2 that the matrix product AB is not equal to BA, in general. One generally says that two matrices do not *permute* or do not *commute*. If $AB = - BA$ then the matrices *anti-commute*.

The elements of a matrix can be real or complex. We take the general case of matrices with complex elements. The transformation formulas (1) will then express the idea that one passes from certain complex variables x_i to other complex variables x_i'. We shall now define some very important special cases of complex matrices.

We say that a matrix is *Hermitian* if the elements that are symmetric with respect to the diagonal are complex conjugates ($a_{ik} = a_{ki}^*$). The diagonal elements of a Hermitian matrix are real. If all of the elements of a Hermitian matrix are real then the matrix will be symmetric with respect to the diagonal. We say that a matrix is *anti-Hermitian* if one has $a_{ik} = -a_{ki}^*$. The diagonal elements of an anti-Hermitian matrix are pure imaginary. The product of two Hermitian matrices is Hermitian only if they commute; if they anti-commute then the product will be anti-Hermitian.

One calls the matrix that is obtained by starting with A and permuting the terms symmetrically with respect to the diagonal and taking complex conjugate quantities the *adjoint matrix to the matrix A* and denotes it by A^+; one will then have $a_{ik}^+ = a_{ki}^*$. One easily proves the formula $(AB)^+ = B^+ A^+$, and it is obvious that $(A^+)^+ = A$.

A matrix is called *diagonal* when only its diagonal elements can be non-zero. A very important diagonal, Hermitian matrix is the identity matrix, which one represents by 1: It is the matrix whose ik element is equal to δ_{ik}.

If one is given a matrix A, and there exists another matrix A^{-1} such that $A \cdot A^{-1} = A^{-1} \cdot A = 1$ then the matrix A^{-1} will be called the *matrix inverse of A*. That matrix inverse (if it exists) is always unique. When A has a finite number of rows and columns, the matrix inverse will always exist if the determinant that is defined with the aid of the table of a_{ik} is not zero. When A has an infinite number of rows and columns, A^{-1} might not exist, and it will be necessary to verify its existence in each special case. One easily verifies the formula $(AB)^{-1} = B^{-1} A^{-1}$.

When A is a matrix with real elements, and one has:

$$\sum_i a_{ij}\, a_{ik} = \delta_{jk}, \tag{3}$$

one says that the matrix A is *orthogonal*. It defines an orthogonal transformation that leaves the quantity $\sum_i x_i^2$ invariant. That is well-known. One can generalize the definition to matrices with complex elements: If A is a complex matrix, and if one has:

$$\sum_i a_{ij}\, a_{ik}^* = \delta_{jk} \tag{4}$$

then one will say that A defines a *complex orthogonal* transformation [*sic*], or furthermore, that it is *unitary*. The quantity $\sum_i x_i\, x_i^*$ remains invariant for a complex orthogonal transformation. The condition (4) can be further written:

$$\sum_i a_{ki}^+ \, a_{ij} = \delta_{kj}, \qquad \text{or} \qquad A^+ A = 1, \qquad A^+ = A^{-1}. \tag{5}$$

Hence, in order for a matrix to be unitary, its adjoint must coincide with its inverse.

Let A once more denote a matrix, and let S be a unitary matrix that has the same number of rows and columns. The matrix $B = S^{-1} A \, S$ is said to be obtained from A by a *canonical transformation.* If A is Hermitian then B will also be Hermitian. Indeed, since, by hypothesis, $S^+ = S^{-1}$ and $A^+ = A$, one will have:

$$B^+ = (S^{-1}AS)^+ = S^+ A^+ (S^{-1})^+ = S^{-1} AS = B. \tag{6}$$

One then has the important theorem: *A canonical transformation transforms a Hermitian matrix into another Hermitian matrix.* ([1]).

2. Matrices in wave mechanics. – Suppose that we know a complete system of normalized, orthogonal functions: $\varphi_1, \ldots, \varphi_n$; we call them *basis functions.* For example, such a system is provided by the set of proper functions of a Hermitian operator.

If one is given a basic system then 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 function that must be developable in terms of the complete sequence $\varphi_1, \ldots, \varphi_n, \ldots$ We then have a relation of the form:

$$A \, (\varphi_i) = \sum_j a_{ji} \, \varphi_j \,, \tag{7}$$

and therefore, by reason of the properties of φ_i :

$$a_{ji} = \int_D \varphi_j^* A(\varphi_i) \, d\tau \,, \tag{8}$$

in which D is the domain of the variables that appear in the φ_i .

By definition, the a_{ji} in formulas (8) are elements of the matrix that is generated by the operator A in the basis system of the φ_i . We also denote that matrix by the letter A; if we would like to make the system of basis functions used more explicit then we might denote it by A^φ.

The matrices thus-obtained can be called the *matrices of wave mechanics.* We shall verify that they indeed satisfy the rules of addition and multiplication of algebraic matrices. In order to do that, consider two linear operators A and B. We will have:

$$A \, (\varphi_i) = \sum_j a_{ji} \varphi_i \,, \qquad B \, (\varphi_i) = \sum_j b_{ji} \varphi_i \,, \tag{9}$$

so

([1])   That statement assumes essentially that the matrix S of the canonical transformation is unitary.

$$(A + B)\,(\varphi_i) = \sum_j (a_{ji} + b_{ji})\,\varphi_j \,. \qquad (10)$$

The element ji of the matrix $A + B$ is then $a_{ji} + b_{ji}$; that is indeed the rule for the addition of algebraic matrices. One has, moreover:

$$AB\,(\varphi_i) = \sum_j b_{ji} A(\varphi_j) = \sum_j b_{ji} \sum_k a_{kj}\,\varphi_k \; = \sum_k \left(\sum_j a_{kj}\,b_{ji} \right) \varphi_k \,. \qquad (11)$$

The element ki of the matrix AB is then $\displaystyle\sum_j a_{kj}\,b_{ji}$; that is indeed the rule for the multiplication of algebraic matrices.

The condition for a matrix A to be Hermitian in wave mechanics is:

$$a_{ji} = \int_D \varphi_j^*\, A(\varphi_i)\,d\tau \; = \; a_{ij}^* \; = \int_D \varphi_j\, A^*(\varphi_i^*)\,d\tau \,. \qquad (12)$$

Now, we know that if that condition is satisfied for all of the functions φ_i then the operator A will be, by definition, Hermitian, and conversely. Hence, the necessary and sufficient condition for a matrix to be Hermitian in wave mechanics is that the operator that it is derived from must also be Hermitian. The Hermiticity is then an intrinsic property of the operator, in the sense that a Hermitian operator will generate a Hermitian matrix in any system of basis functions.

Since all of the operators that we shall consider in wave mechanics are Hermitian, the matrices that correspond to them will likewise be Hermitian.

3. Mean values in wave mechanics. – Imagine a corpuscle, and assume that the wave function Ψ that it is associated with is known. On the other hand, let one of the mechanical quantities that are attached to the corpuscle correspond to the operator A in the new mechanics. When speaking of that operator, we shall say "the quantity A," to abbreviate.

The general principles that were stated in the last chapter permit us to predict the possible magnitudes of the quantity A and their respective probabilities. Since there are, in general, several possible values with non-zero probabilities, one cannot speak unequivocally of the value of the quantity A at each instant, but one can speak of its mean value, where that mean value is defined in the usual fashion as the sum of the products of each possible value with the corresponding probability. If α_i and φ_i denote the proper values and proper functions of the operator A, and if the wave function Ψ admits the development:

$$\Psi = \sum_i c_i\,\varphi_i \qquad (13)$$

then the mean value $\overline{A}$ will be, from general principles:

$$\overline{A} = \sum_i \alpha_i \,|\, c_i\,|^2 \,. \tag{14}$$

That can be written in the form that essential in wave mechanics:

$$\overline{A} = \int_D \Psi^* A(\Psi)\, d\tau \,. \tag{15}$$

The equivalence of (14) and (15) results from the formula:

$$\int_D \Psi^* A(\Psi)\, d\tau = \int_D \sum_i c_i^* \varphi_i^* \cdot A\left(\sum_k c_k \,\varphi_k\right) d\tau = \sum_{ik} c_i^* c_k \alpha_k \int_D \varphi_i^* \varphi_k \, d\tau \,, \tag{16}$$

and from the fact that the functions φ_i are normalized and orthogonal. The mean value $\overline{A}$ that is defined by (14) and (15) is obviously a real quantity.

The reasoning by which we just established the fundamental formula (15) is rigorously valid only for complete operators A with no continuous spectrum and no multiple proper values. However, it is easy to extend it to incomplete or degenerate operators and continuous spectra: Formula (15) is general.

The form of the expression (15) for $\overline{A}$ will permit us to say that $\Psi^* A\,(\Psi)$ is the *density of the mean value* of the quantity A. However, that "density" has a very different nature from the ones that one considers in classical theories. Indeed, the integration element $\Psi^* A\,(\Psi)\, d\tau$ (which is generally complex) cannot at all be considered to be a certain amount of the quantity A that is localized in the element $d\tau$, and only the integral (15), which is always real, will have any physical sense. That is an important remark that one must keep in mind.

Formula (15) provides a statistical interpretation for the matrices of wave mechanics, or at least, their diagonal elements. Once more, we shall show that by arguing on the basis of complete, non-degenerate operators with no continuous spectrum, since the general argument would involve some complications that would not alter the result. Suppose that the development of the wave function Ψ in proper functions φ_i of the operator A reduces to just one term. One will then have:

$$\Psi = c_i \,\varphi_i \,, \tag{17}$$

with $|\, c_i \,| = 1$, since Ψ is always assumed to be normalized. In that case, we will be sure that a measurement of the quantity A will yield the value α_i. Now, let another quantity B be attached to the corpuscle and correspond to an operator B. We will then obtain the mean value of the quantity B by applying the formula (15), and we will find:

$$\overline{B} = \int_D \Psi^* B(\Psi)\, d\tau = \int_D \varphi_i^* B(\varphi_i)\, d\tau \,. \tag{18}$$

Now, the second integral in (18) is nothing but the diagonal element of the index *ii* of the matrix that is generated by the operator A in the system of functions φ_i. One then has the theorem:

The diagonal element with index ii in the matrix that is generated by the operator B in the system of proper functions of the operator A is equal to the mean value of the quantity B **when one knows that the quantity** *A* **has the value** α_i .

4. Mean value of a coordinate. Ehrenfest's theorem. – Consider one of the coordinates of a corpuscle – for example, the *x* coordinate. From formula (15), and in accord with the interference principle, its mean value is:

$$\overline{x} = \int_D x \, \Psi^* \Psi \, d\tau . \tag{19}$$

That is then the *x* coordinate of a *fictitious fluid* whose density at each point will be given by:

$$\rho = \Psi \Psi^*. \tag{20}$$

We call that fictitious fluid the *probability fluid*. The quantity of that fluid that is contained in a volume element $d\tau$ is $\Psi \Psi^* \, d\tau$ ([1]), the total quantity of fluid will remain constant in time, and that will be equal to 1 if the function Ψ has been normalized. Since the distribution of the probability fluid varies in the course of time, we shall attribute a velocity with it at each point in the course of time, which we define by the formula:

$$\mathbf{u} = \frac{1}{\Psi \Psi^*} \frac{h}{4\pi i m} [\Psi \text{ grad } \Psi^* - \Psi^* \text{ grad } \Psi]. \tag{21}$$

It is easy to show that under those conditions, the probability fluid will be preserved in the course of time, and that proof will incidentally establish the following proposition, whose exactitude we have assumed: *If the function* Ψ *is normalized at a given instant then it will remain normalized at any other point in time.*

Here is the proof in question: One easily infers from the wave equation that the function Ψ obeys [viz., equation (5) of the preceding chapter], along with its conjugate equation, that:

$$\frac{\partial}{\partial t}(\Psi \Psi^*) = (\Psi^* \Delta \Psi - \Psi \Delta \Psi^*) = -\frac{h}{4\pi i m} \sum_{x,y,z} \left[\frac{\partial}{\partial x} \left(\Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x} \right) \right], \tag{22}$$

or again, from the definitions (20) and (21):

$$\frac{\partial \rho}{\partial t} + \text{div} \, (\rho \, \mathbf{u}) = 0. \tag{23}$$

([1]) From the interference principle, that quantity is then equal to the "probability of presence" of the corpuscle in the element $d\tau$.

Now, equation (23) is the hydrodynamical continuity equation that expresses precisely the conservation of the fluid whose distribution and motion are defined by (20) and (21).

We now make a remark. From our definition of a mean value, if $f(x, y, z, t)$ is a certain scalar or vectorial function then we shall call the quantity:

$$\overline{f}(t) = \int_D \Psi^* \cdot f(x, y, z, t) \cdot \Psi \, d\tau \tag{24}$$

the *mean value of that function at the instant t.*

That being the case, we can state an important theorem that is due to Ehrenfest:

THEOREM: *The point whose coordinates are $\overline{x}$, $\overline{y}$, $\overline{z}$ displaces in the course of time as it would under the laws of classical mechanics for a material point of mass m under the action of a force that is equal to the mean value $\overline{f}(t)$ of the real force at each instant t.*

From (19), and upon employing (22) and integrating by parts, one will first get:

$$\frac{d\overline{x}}{dt} = \int_D x \frac{\partial(\Psi\Psi^*)}{\partial t} d\tau = -\frac{h}{4\pi im} \int_D x \sum_{x,y,z} \frac{\partial}{\partial x} \left[\Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x} \right] d\tau$$

$$= \frac{h}{4\pi im} \int_D x \left(\Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x} \right) d\tau = \int_D u_x \Psi^* \Psi \, d\tau = \overline{u}_x. \tag{25}$$

One will then find that:

$$\frac{d^2\overline{x}}{dt^2} = \frac{h}{4\pi im} \int_D \left[\frac{\partial\Psi}{\partial t} \frac{\partial\Psi^*}{\partial x} - \frac{\partial\Psi}{\partial x} \frac{\partial\Psi^*}{\partial t} + \Psi \frac{\partial^2\Psi^*}{\partial x \partial t} - \Psi^* \frac{\partial^2\Psi}{\partial x \partial t} \right] d\tau$$

$$= \frac{h}{2\pi im} \int_D \left(\frac{\partial\Psi}{\partial t} \frac{\partial\Psi^*}{\partial x} - \frac{\partial\Psi^*}{\partial x} \frac{\partial\Psi}{\partial t} \right) d\tau, \tag{26}$$

which will give, by virtue of the equation of propagation:

$$\frac{d^2\overline{x}}{dt^2} = -\frac{h}{8\pi^2 m^2} \int_D \left[\frac{\partial\Psi^*}{\partial x} \left(\Delta\Psi - \frac{8\pi^2 m}{h^2} U\Psi \right) + \frac{\partial\Psi}{\partial x} \left(\Delta\Psi^* - \frac{8\pi^2 m}{h^2} U\Psi^* \right) \right] d\tau. \tag{27}$$

Now, Green's theorem gives (after integrating by parts):

$$\int_D \left(\frac{\partial\Psi^*}{\partial x} \Delta\Psi + \frac{\partial\Psi}{\partial x} \Delta\Psi^* \right) d\tau = \int_D \left[\frac{\partial\Psi^*}{\partial x} \Delta\Psi - \Psi^* \frac{\partial}{\partial x}(\Delta\Psi) \right] d\tau = 0, \tag{28}$$

in such a way that (27) will give us:

$$m\frac{d^2\overline{x}}{dt^2} = \int_D U \frac{\partial}{\partial x}(\Psi\Psi^*)\,d\tau = \int_D \left(-\frac{\partial U}{\partial x}\right)\Psi\Psi^*\,d\tau = -\frac{\partial U}{\partial x} = \overline{f_x}. \tag{29}$$

That is indeed the expression for Ehrenfest's theorem for $\overline{x}$, and one will likewise establish corresponding formulas for $\overline{y}$ and $\overline{z}$.

In order to conclude this paragraph, we shall once more give the definitions of the mean densities of electric charge and current that correspond to a corpuscle of charge ε when one knows the function Ψ. Since the charge ε is a physical quantity that is capable of taking just one value, it will be equal to its mean value, and one can write:

$$\overline{\varepsilon} = \varepsilon = \int_D \varepsilon\,\Psi\Psi^*\,d\tau. \tag{30}$$

One can then consider the quantity:

$$\delta = \varepsilon\,\Psi\Psi^* \tag{31}$$

to be the mean density of the electric charge that is associated with the corpuscle. On the other hand, from the classical viewpoint, the corpuscle of charge ε that moves with a velocity of $\mathbf{v}$ is equivalent to a current element $\boldsymbol{i} = \varepsilon\,\mathbf{v}$. Here, we have to replace v_x with 1 / $m\,p_x$, etc., and the three components i_x, i_y, i_z of the current in question will correspond to the operators:

$$-\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial x}, \quad -\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial y}, \quad -\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial z},$$

resp.
From (15), the mean value of i_x, for example, will always be:

$$\overline{i_x} = -\frac{h\varepsilon}{2\pi im}\int_D \Psi^*\frac{\partial\Psi}{\partial x}\,d\tau = \frac{h\varepsilon}{4\pi im}\int_D \left(\Psi\frac{\partial\Psi^*}{\partial x} - \Psi^*\frac{\partial\Psi}{\partial x}\right)d\tau, \tag{32}$$

and one will have analogous expressions for $\overline{i_y}$ and $\overline{i_z}$. It results from those expressions that the vectorial quantity:

$$\boldsymbol{i} = \frac{h\varepsilon}{4\pi im}[\Psi\,\mathrm{grad}\,\Psi^* - \Psi^*\,\mathrm{grad}\,\Psi] \tag{33}$$

can be considered to be the mean density of the electric current that is associated with the particle. Upon comparing formulas (30) and (33) with formulas (20) and (21), one will see that the densities δ and $\boldsymbol{j}$ are the charge and current densities that exist, from classical theory, if the charge ε of the corpuscle is distributed throughout the probability fluid in proportion to its density $\Psi\Psi^*$.

5. First integrals in wave mechanics. – In classical mechanics, one calls a mechanical quantity that is expressed in terms of the coordinates, momenta, and

sometimes time, a *first integral* when it remains constant in the course of motion by virtue of the equations of motion themselves.

How does one define a first integral in the new mechanics? Here is the answer that one must give to that question: If a mechanical quantity corresponds to an operator *A* then that quantity will be a first integral for a well-defined problem (i.e., for a given form to the Hamiltonian *H*) if one has:

$$\frac{\partial A}{\partial t} + \frac{2\pi i}{h}(AH - HA) = 0, \tag{34}$$

in which $\partial A / \partial t$ represents the operator that one obtains by formally differentiating the expression for *A* with respect to the variable *t*. If *A* does not depend upon *t* then $\partial A / \partial t$ will be zero, and the condition (34) will express simply the idea that *A* commutes with *H*. One shows moreover that if *H* is independent of time then the condition (34) will have the following significance: The matrix elements that are generated by the operator *A* in the system of proper functions of the operator *H* are constant.

If the field is constant then *H* will not depend upon time, so energy will obviously be a first integral: One the recovers a classical theorem. If the *x* component of the field is zero then *H* will not depend upon *x*, and it will commute with p_x : The *x*-component of the quantity of motion will be a first integral, as in classical mechanics, etc.

The most interesting case for us will be that of angular momentum. When the field presents cylindrical symmetry around an axis *oz*, *H* will not depend upon the azimuth φ around that axis. Upon taking direct axes and agreeing that a rotation in the *xy*-plane in the positive sense will correspond to a positive rotation, the operator that corresponds to the angular momentum around *oz* will be:

$$M_z = x\, p_y - y\, p_x = \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right). \tag{35}$$

Upon taking spherical coordinates whose polar axis is *oz*, one will find that:

$$M_z = -\frac{h}{2\pi i}\frac{\partial}{\partial \varphi}. \tag{36}$$

As a result, M_z will be independent of *t*, so it will commute with *H*, and it will be a first integral. If the field has spherical symmetry around 0 then each of the angular momenta M_x, M_y, M_z will be first integrals.

We can simplify the form of the condition (34) by introducing the operator:

$$L = H - \frac{h}{2\pi i}\frac{\partial}{\partial t}. \tag{37}$$

Since one has:

$$\frac{\partial}{\partial t} \cdot A\,(f) = \frac{\partial A}{\partial t}(f) + A\left(\frac{\partial f}{\partial t} \right),$$

if f is an arbitrary function then the operator $\partial A / \partial t$ will be equivalent to the operator $\dfrac{\partial}{\partial t} \cdot$

$A - A \cdot \dfrac{\partial}{\partial t}$. The condition (34) can then be written:

$$0 = (AH - HA) + \frac{h}{2\pi i}\left(\frac{\partial}{\partial t}\cdot A - A\cdot\frac{\partial}{\partial t}\right) = A\left(H - \frac{h}{2\pi i}\frac{\partial}{\partial t}\right) - \left(H - \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)A, \tag{38}$$

or simply:

$$AL - LA = 0. \tag{39}$$

The condition for the quantity A to be a first integral is then simply that the operator A must commute with L.

_____

RELATIVISTIC FORM OF WAVE MECHANICS WITH ONE WAVE FUNCTION

——

1. Review of some formulas from relativistic mechanics

Long before the appearance of the new mechanics, the introduction of the principle of relativity into mechanics led Einstein to modify the classical equations of Newtonian dynamics. That modification implied only a simple change of form in certain formulas, moreover. Einstein's relativistic mechanics preserved all of the classical concepts of a material point, velocity, trajectory, mechanical determinism, etc. In comparison to the new mechanics, the Einstein's dynamics then appeared to be only a slight modification of the classical theory that had the goal of making it conform to the principle of relativity. In paragraph 2 of Chapter V, we started with the formulas of Newtonian mechanics in order to obtain the equation the equation of propagation of wave mechanics. We thus obtained a wave mechanics that was naturally non-relativistic. In order to get relativistic wave mechanics, it seems quite natural to operate as in paragraph 2 of that chapter, but while starting with formulas from Einstein's theory. In order to do that, we begin by recalling some of those formulas.

In relativistic mechanics, every corpuscle is characterized by an invariant quantity m_0, namely, its proper mass. One of the fundamental principles of the theory of relativity is the proportionality of mass and energy, so a corpuscle of mass m_0 will possess an internal energy or "proper energy," even at rest, and it will be given by:

$$W_0 = m_0 \, c^2, \tag{1}$$

in which c is the speed of light *in vacuo*. If the corpuscle moves with a velocity of $v = \beta c$ then its energy will be:

$$W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} \, . \tag{2}$$

There will still be a proportionality between energy and mass, on the condition that one must consider the mass to have increased as a result of the motion, and to have become $m_0 / \sqrt{1 - \beta^2}$.

One can call the quantity:

$$T = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} - m_0 \, c^2 = m_0 \, c^2 \left(\frac{1}{\sqrt{1 - \beta^2}} - 1 \right), \tag{3}$$

the *kinetic energy* of the corpuscle that has been animated with a speed of βc, and it will represent the increase of the energy due to the motion.

The preceding formulas are valid in the absence of a field. If the corpuscle is subject to a force field that is derived from the potential function then one must write:

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + U(x, y, z, t), \tag{4}$$

instead of (2).

If β is small compared to unity (i.e., $v \ll c$) then one will get, in the first approximation:

$$W = m_0 c^2 + \tfrac{1}{2} m_0 v^2 + U(x, y, z, t), \tag{5}$$

and if one sets:

$$E = W - m_0 c^2 \tag{6}$$

then one will recover the classical formula. E is then the energy in the sense of classical mechanics, which differs from the energy W of relativistic mechanics by the constant term of proper energy $m_0 c^2$.

In Einstein's theory, the quantity of motion of the corpuscle of mass m_0 that is animated with the speed $v = \beta c$ is:

$$\mathbf{p} = \frac{m_0}{\sqrt{1-\beta^2}} \mathbf{v} . \tag{7}$$

It is, in summary, equal to the product of the velocity with the mass in motion $m_0 / \sqrt{1-\beta^2}$.

The three components of the quantity of motion and the quantity:

$$\frac{m_0 c}{\sqrt{1-\beta^2}} \qquad (= W / c, \text{ when there is no field})$$

constitute the four components of a space-time vector.

The preceding formulas must be modified in the very important case of a corpuscle of electric charge ε that displaces in an electromagnetic field. One knows that the electric field $\mathbf{h}$ and the magnetic field $\mathbf{H}$ can be defined by the relations:

$$\mathbf{H} = \operatorname{rot} \mathbf{A}, \qquad \mathbf{h} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{8}$$

with the aid of a scalar potential $V(x, y, z, t)$ and a vector potential $\mathbf{A}(x, y, z, t)$.

The force that the field exerts upon the corpuscle of charge ε that is animated with a velocity of $\mathbf{v}$ will then be given by the Lorentz formula:

$$\mathbf{f} = \varepsilon \left(\mathbf{h} + \frac{1}{c} [\mathbf{v} \cdot \mathbf{H}] \right) . \tag{9}$$

The energy of the corpuscle is then:

$$W = \frac{m_0 c}{\sqrt{1 - \beta^2}} + \varepsilon V (x, y, z, t), \tag{10}$$

and its quantity of motion is:

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \beta^2}} + \varepsilon \frac{\mathbf{A}(x, y, z, t)}{c} . \tag{11}$$

The new fact here is that the quantity of motion, just like the energy, contains a term that depends upon the field, in addition to the term that depends upon the velocity. In general, the quantity of motion $\mathbf{p}$ will no longer be directed along the velocity. The components of $\mathbf{p}$ and the quantity W / c will then once more define the components of a vector in space-time.

Equation (10) and (11) give us the relation:

$$\frac{1}{c^2} (W - \varepsilon V)^2 - \sum_{x,y,z} \left(p_x - \frac{\varepsilon A_x}{c} \right)^2 - m_0^2 c^2 = 0, \tag{12}$$

which must play an essential role when one wants to get a relativistic wave equation in wave mechanics.

If we solve (12) for W then we will get:

$$W = c \sqrt{m_0^2 c^2 + \sum_{x,y,z} \left(p_x - \frac{\varepsilon A_x}{c} \right)^2} + \varepsilon V. \tag{13}$$

The right-hand side of that equation can be denoted by $H (x, y, z, p_x, p_y, p_z, t)$, and it is the relativistic Hamiltonian function; however, it is not a rational function.

2. Relativistic wave mechanics. – In order to obtain the wave equation of relativistic wave mechanics, it seems completely natural to proceed as in paragraph 2 of Chapter V by starting, no longer with the formulas of classical dynamics, but with those of relativistic dynamics. Unfortunately, that immediately presents a complication: Since the Hamiltonian function that is defined by the right-hand side of (13) is not rational, the expressions that one will obtain by replacing p_x with $-\dfrac{h}{2\pi i} \dfrac{\partial}{\partial x}$, etc., will not be rational in $\partial / \partial x$, etc., and will not represent well-defined operators. One cannot apply the method

of paragraph 2, Chapter V literally then; i.e., one cannot pose the wave equation in the form $H(\Psi) = \dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$.

Meanwhile, there is an indirect means of avoiding the difficulty by using equation (12), instead of (13). In order to do that, one remarks that already in the non-relativistic method of Chapter V, when one passes from the classical equation $H = E$ to the wave equation $H(\Psi) = \dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$, one replaces, in short, the energy E with the operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$. It would then seem natural to replace W with that operator here, and that would be much more logical from the relativistic viewpoint, since energy and the quantity of motion define a space-time vector, so if one replaces each p_i with $-\dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_i}$ then W must be replaced with $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ ([1]).

Upon making those substitutions in the left-hand side of (12), one will get:

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} - \varepsilon V\right)^2 - \sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{\varepsilon}{c}A_x\right)^2 - m_0^2 c^2 \; ; \tag{14}$$

i.e., a rational operator.

Upon applying the operator (14) to the function Ψ and upon equating it to zero, one will find that:

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} - \varepsilon V\right)^2 \Psi - \sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{\varepsilon}{c}A_x\right)^2 \Psi = m_0^2 c^2 \Psi, \tag{15}$$

and that equation can be regarded as the natural relativistic extension of the wave equation of the original wave mechanics.

If we develop equation (15), while taking into account the Lorentz relation between the potentials:

$$\frac{1}{c}\frac{\partial V}{\partial t} + \operatorname{div}\mathbf{A} = 0 \tag{16}$$

then we will obtain:

$$\frac{1}{c^2}\frac{\partial^2 \Psi}{\partial t^2} - \Delta\Psi - \frac{4\pi i}{m}\frac{\varepsilon V}{c^2}\frac{\partial \Psi}{\partial t} - \frac{4\pi i}{m}\sum_{x,y,z}\frac{\varepsilon A_x}{c^2}\frac{\partial \Psi}{\partial x} + \frac{4\pi^2}{h^2}\left[m_0^2 + \frac{\varepsilon^2}{c^2}(V^2 - \mathbf{A}^2)\right]\Psi = 0. \tag{17}$$

([1]) The difference in sign is explained by remarking that:

$$\frac{\partial}{\partial x}, \; \frac{\partial}{\partial y}, \; \frac{\partial}{\partial z}, \; \frac{\partial}{\partial t}$$

are covariant components, whereas p_x, p_y, p_z, W are the contravariant components of the world-impulse.

For reasons that we shall explain later, Dirac considered that relativistic equation of propagation to be insufficient. Moreover, we see immediately that it differs greatly from the non-relativistic equation in a very important respect: It is a differential equation that has second order in time, instead of one that has first order in it.

In the important case in which the electromagnetic field is zero ($V = \mathbf{A} = 0$), equation (17) can be written:

$$\Delta\Psi - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} = \frac{4\pi^2}{h^2}m_0^2 c^2\Psi , \tag{18}$$

and it will admit the monochromatic plane wave:

$$\Psi = C\, e^{\frac{2\pi i}{h}[Wt - p_x x - p_y y - p_z z]} \tag{19}$$

as a particular solution, because, from (12), one will then have $W^2 / c^2 - p_x^2 - p_y^2 - p_z^2 = m_0^2 c^2$. The wave (19) possesses a frequency $v = W / h$, and a wave length $\lambda = h / p$. It corresponds to the uniform, rectilinear motion of a corpuscle whose energy and quantity of motion are perfectly determined, while its position is entirely indeterminate. If one compares formula (19) with formula (7) in Chapter V then one will see that the only essential difference that is introduced by relativity here is the substitution of the complete energy W for the energy $E = W - m_0 c^2$ of classical mechanics.

3. Probability density and current that correspond to equation (17). – The application of the general principles that were stated in Chapters V and VI to the wave equation (17) raises some great difficulties. In particular, there will no longer be any clear meaning to saying that the wave function must be normalized, because one can no longer show with equation (17) that if Ψ is normalized at a given instant then it will always remain normalized. As we will see in Chapter X, that difficulty is essentially attached to the fact that equation (17) has second order in time, so its solution will not be determined when one knows only the initial form of the function Ψ.

It is, nonetheless, still possible with equation (17) to consider a probability fluid that is conserved in the course of time by virtue of the equation of propagation itself, but the density of that fluid will be expressed as a function of not just Ψ, but also $\partial\Psi / \partial t$. Indeed, set:

$$\rho = -\frac{h}{4\pi i\, m_0 c^2}\left(\Psi\frac{\partial\Psi^*}{\partial t} - \Psi^*\frac{\partial\Psi}{\partial t}\right) - \frac{\varepsilon}{m_0 c^2}V\,\Psi\Psi^* , \tag{20}$$

$$\rho\,\mathbf{u} = \frac{h}{4\pi i\, m_0 c^2}\left(\Psi\,\mathrm{grad}\,\Psi^* - \Psi^*\,\mathrm{grad}\,\Psi\right) - \frac{\varepsilon}{m_0 c^2}\mathbf{A}\,\Psi\Psi^* , \tag{20$'$}$$

which are equations that will define the distribution and the motion of a fictitious probability fluid, since ρ and $\rho\,\mathbf{u}$ are real.

Now, write the conjugate equation to (17):

$$\frac{1}{c^2}\frac{\partial^2 \Psi^*}{\partial t^2} - \Delta\Psi^* + \frac{4\pi i}{h}\frac{\varepsilon}{c^2}V\frac{\partial\Psi^*}{\partial t} + \frac{4\pi i}{h}\frac{\varepsilon}{c}\sum_{x,y,z}A_x\frac{\partial\Psi^*}{\partial x} + \frac{4\pi^2}{h^2}\left[m_0^2 c^2 + \frac{\varepsilon^2}{c^2}(V^2 - \mathbf{A}^2)\right]\Psi^*$$

$$= 0. \qquad (17^*)$$

Multiply (17) by Ψ^* and subtract (17*), when it is multiplied by Ψ. On the one hand, one will find the terms:

$$\frac{1}{c^2}\left(\Psi^*\frac{\partial^2 \Psi}{\partial t^2} - \Psi\frac{\partial^2 \Psi^*}{\partial t^2}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2}V\left(\Psi\frac{\partial\Psi^*}{\partial t} + \Psi^*\frac{\partial\Psi}{\partial t}\right),$$

which one can write as:

$$\frac{\partial}{\partial t}\left[\frac{1}{c^2}\left(\Psi^*\frac{\partial^2 \Psi}{\partial t^2} - \Psi\frac{\partial^2 \Psi^*}{\partial t^2}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2}V\,\Psi\Psi^*\right] + \frac{4\pi i}{h}\frac{\varepsilon}{c^2}\Psi\Psi^*\frac{\partial V}{\partial t},$$

and on the other hand, one has terms like:

$$-\Psi^*\Delta\Psi + \Psi\Delta\Psi^* - \frac{4\pi i}{h}\frac{\varepsilon}{c}\sum_{x,y,z}\left(A_x\Psi^*\frac{\partial\Psi}{\partial x} + A_x\Psi\frac{\partial\Psi^*}{\partial x}\right),$$

which one can write as:

$$\sum_{x,y,z}\left\{\frac{\partial}{\partial x}\left[\left(\Psi\frac{\partial\Psi^*}{\partial x} - \Psi^*\frac{\partial\Psi}{\partial x}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c}A_x\Psi\Psi^*\right] + \frac{4\pi i}{h}\frac{\varepsilon}{c}\Psi\Psi^*\frac{\partial A_x}{\partial x}\right\}.$$

When summed, we will get the equation:

$$\frac{\partial}{\partial t}\left[\frac{1}{c^2}\left(\Psi^*\frac{\partial^2 \Psi}{\partial t^2} - \Psi\frac{\partial^2 \Psi^*}{\partial t^2}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2}V\,\Psi\Psi^*\right]$$

$$+ \operatorname{div}\left[\Psi\,\operatorname{grad}\Psi^* - \Psi^*\,\operatorname{grad}\Psi - \frac{4\pi i}{h}\frac{\varepsilon}{c}\mathbf{A}\,\Psi\Psi^*\right]$$

$$+ \frac{4\pi i}{h}\frac{\varepsilon}{c^2}\Psi\Psi^*\left(\frac{1}{c}\frac{\partial V}{\partial t} + \operatorname{div}\mathbf{A}\right) = 0. \qquad (21)$$

The last term is zero, by reason of the Lorentz relation (16). If one multiplies it by $\dfrac{h}{4\pi i\,m_0}$ and takes (20) and (20′) into account then one will get:

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\,\mathbf{u}) = 0; \qquad (22)$$

i.e., the equation that expresses the conservation of the fluid ([1]).

We can make the distribution and motion of the probability fluid that is associated with a corpuscle with an electric charge ε correspond to an electric charge density $\rho\,\varepsilon$ and a current density $\rho\,\varepsilon\,\mathbf{u}$, as we did at the end of the preceding chapter in the non-relativistic case.

One can account for some difficulties that arise when one attempts to reconcile the general principles that were assumed in Chapter V with the relativistic form (17) of the wave equation by referring to the manner by which the interference principle was justified in paragraph 5 of Chapter V. It will result from the argument that was presented in that regard that the probability for a measurement that is made at the instant t to permit one to localize the corpuscle into an element $d\tau$ of space will be $\Psi\Psi^{*}\,d\tau$. That result is obtained without making any hypothesis about the wave equation that the wave function Ψ must obey, so it must still be valid here, as well. Meanwhile, the density ρ of the probability fluid that is given by the relation (20), which is a density whose expression is imposed by the necessity of satisfying the conservation condition, does not reduce to $\Psi\Psi^{*}$. The general principles of Chapter V then contradict the relativistic form (17) for the equation of propagation. It was the desire to avoid such a contradiction that led Dirac to found the relativistic wave mechanics of a corpuscle upon a different basis than equation (17).

————————

_____

([1]) It would seem that one could define ρ and $\rho\,\mathbf{u}$ by taking the right-hand sides of (20) and (20′) and multiplying them by an arbitrary real constant, but the definitions (20) and (20′) are imposed when one desires to recover the expression $\rho = \Psi\Psi^{*}$ in the Newtonian approximation, as one will easily see.

SUCCESSES AND FAILURES OF WAVE MECHANICS WITH ONE WAVE FUNCTION

―――

1. Calculation of quantized energies. Example of the hydrogen atom.

As an application of its general principles, (non-relativistic) wave mechanics determines the stationary states of quantized systems by calculating the proper values of the corresponding Hamiltonian operator. That new method of quantization, which was inaugurated by the celebrated work of Schrödinger ([1]), led one to recover the results of the old quantum theory in certain cases and to then correct the old results in a sense that would better conform to experiments (e.g., the linear oscillator) in other cases.

Here, we shall recall only the quantization of the hydrogen atom, and like Schrödinger we shall appeal to the non-relativistic theory.

We consider an electron of mass m and charge $-e$ in a field that is produced by a fixed nucleus of charge $+e$.

Write the equation:

$$H(a) = E \cdot a, \tag{1}$$

upon taking H to be the operator $\dfrac{1}{2m}[p_x^2 + p_y^2 + p_z^2] - \dfrac{e^2}{r}$.

We then get:

$$\Delta a + \frac{8\pi^2 m}{h^2}\left[E + \frac{e^2}{r}\right] a = 0. \tag{2}$$

Take polar coordinates r, θ, φ around the nucleus and set:

$$a(r, \theta, \varphi) = R(r)\, Y(\theta, \varphi). \tag{3}$$

Upon taking the form of the Laplacian in polar coordinates into account, we will have:

$$Y\frac{d^2 R}{dr^2} + 2\frac{Y}{r}\frac{dR}{dr} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{8\pi^2 m}{h^2}\left[E + \frac{e^2}{r}\right]RY = 0, \tag{4}$$

or furthermore:

$$r^2\left[\frac{1}{R}\frac{d^2 R}{dr^2} + \frac{1}{r}\frac{1}{R}\frac{dR}{dr} + \frac{8\pi^2 m}{h}\left(E + \frac{e^2}{r}\right)\right] = -\frac{1}{Y}\left[\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)\right]. \tag{5}$$

―――――――――

([1]) See E. SCHRÖDINGER, *Abhandlungen zur Wellenmechanik*, J. A. Barth, Leipzig.

One of the two sides of (5) depends upon only the radius vector, while the other depends upon only the polar angles, so they must both be equal to the same constant λ, and one will have:

$$\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \lambda\,Y = 0. \tag{6}$$

One shows that equation (6) has solutions that are finite, uniform, and continuous on the sphere of unit radius only if one has:

$$\lambda = l\,(l+1), \quad \text{with} \quad l = 0, 1, 2, \dots \tag{7}$$

The values (7) of λ are proper values of equation (6). A proper value that is defined by a certain integer value for l corresponds to $(2l+1)$ proper functions (un-normalized):

$$Y_l^m(\theta,\varphi) = e^{im\varphi}\,P_l^m(\cos\theta) = e^{im\varphi}\sin^m\theta\,\frac{d^{l+m}}{(d\cos\theta)^{l+m}}\cdot(1-\cos^2\theta)^l. \tag{8}$$

The functions Y_l^m are Laplace's spherical functions. They define a complete system for the variables φ and θ, which justifies the decomposition (3) *a posteriori*.

Having posed that, it will again result from (5) that R must satisfy the equation:

$$\frac{d^2 R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + R\left(A + \frac{B}{r} + \frac{C}{r^2}\right) = 0, \tag{9}$$

with:

$$A = \frac{8\pi^2 mE}{h^2}, \quad B = \frac{8\pi^2 m}{h^2}e^2, \quad C = l\,(l+1). \tag{10}$$

Schrödinger proved that all of the positive values of E are proper values of (9) and form a continuous spectrum, in turn. However, as in classical mechanics, those positive proper values for the energy correspond to free motions of the electron outside the atom, and are of no interest to us here.

In order to find the negative proper values, we introduce the real variable:

$$\rho = 2\sqrt{-A}\,r = \frac{4\pi}{h}\sqrt{-2mE}\,\cdot r. \tag{11}$$

It is obvious in equation (9) that for very large r, R will have the asymptotic form $e^{-\rho/2}$, so the solution $e^{+\rho/2}$ must be discarded, because it will be infinite at infinity. Therefore, set:

$$R = e^{-\frac{\rho}{2}}\,\upsilon\,(\rho). \tag{12}$$

Upon substituting (11) and (12) into (9), we easily find:

$$\frac{d^2v}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{dv}{d\rho} + \left[\left(\frac{B}{2\sqrt{-A}} - 1\right)\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right]v = 0. \tag{13}$$

That is a linear differential equation that admits only the singular point $\rho = 0$ at a finite distance. The theory of linear equations then permits one to easily see that equation (13) admits just one solution that is regular in a neighborhood of the point $\rho = 0$ and finite at that point, and that solution will have the form:

$$v(\rho) = \sum_v a_v \,\rho^{l+v} . \tag{14}$$

Upon substituting (14) in (13), one will find the recurrence relation:

$$[(v + l + 1)(v + 1) + 2(v + l + 1) - l(l + 1)]\, a_{v+1} = \left[v + l + 1 - \frac{B}{2\sqrt{-A}}\right] a_v . \tag{15}$$

The function $R(\rho)$ will be zero at infinity if all of the a_v are zero when one starts with a certain a_p. From formula (15), in order for that to be true, it will be necessary that one must have:

$$\frac{B}{2\sqrt{-A}} = p + l + 1 = n \qquad (n = \text{integer} \geq 1). \tag{16}$$

Hence, from the values (10):

$$E_n = - \frac{2\pi^2 m e^4}{n^2 h^2} . \tag{17}$$

Non-relativistic wave mechanics then gives the Bohr formula.

We remark that there is a pronounced degeneracy here, because, from (16), a proper value E_n (i.e., a given value of n) will correspond to n possible values of l, namely, 0, 1, …, $n - 1$, and each value of l will correspond to $2l + 1$ spherical functions Y_l^m. Hence, a proper value E_n will correspond to some proper functions $a = RY$, and their total number will be equal to $\sum_{l=0}^{n-1}(2l+1) = 2\,\frac{n(n-1)}{2} + n = n^2$. All of the proper values will then be multiple, except for the one that corresponds to $n = 1$.

The quantum number l corresponds to the number $k - 1$ of the old quantum theory. It can taken on the values 0, 1, …, $n - 1$, while the number k in the old theory could take on the values 1, 2, …, n. The number m depends upon the choice of polar axis, which is entirely arbitrary in the absence of an external field. It is easy to verify that the angular momenta M_x, M_y, M_z are first integrals, since that would follow from the spherical symmetry of the Coulomb field.

We finally note that upon repeating the calculations for an atom with a central charge $+ Ne$ that has been ionized $(N - 1)$ times, we will get:

$$E_n = -\frac{2\pi^2 m e^4 N^2}{n^2 h^2}.$$

(18)

Hence, one can infer an approximate justification for Moseley's law, as in the old theory.

2. Fine structure and relativistic wave mechanics. – We have recovered Bohr's simple result by calculating the possible values for the energy of the hydrogen atom. It is quite natural for us to now demand to know whether we can recover Sommerfeld's fine structure by employing the relativistic form of wave mechanics that was developed in the preceding chapter. However, that immediately presents a complication, since we cannot seek to determine the proper values of the Hamiltonian operator, because that operator was not well-defined, as we saw in the theory of the last chapter. Nevertheless, there is a very natural way of avoiding that difficulty (although it does not agree with the general principles). Indeed, seeking the proper values of energy amounts to looking for proper frequencies of the wave equation. Therefore, take the relativistic wave equation (17) of the last chapter and suppose that Ψ is a monochromatic wave that depends upon time only by way of the factor $e^{\frac{2\pi i}{h}Wt}$. Upon noting that $\mathbf{A} = 0$ for the hydrogen atom, one will then find that:

$$\Delta\Psi + \frac{4\pi^2}{h^2 c^2}[(W - \varepsilon V)^2 - m_0^2 c^2]\ \Psi = 0.$$

(19)

In the case of the hydrogen atom, $\varepsilon = -e$ and $V = e / r$. Inside the bracket in (19), one can write:

$$\left(m_0 c^2 + E + \frac{e^2}{r}\right)^2 - m_0^2 c^4 = E^2 + 2m_0\, c^2\, E + \frac{2e^2}{r}(m_0 c^2 + E) + \frac{e^4}{r^2},$$

(20)

and since, by hypothesis, $\Psi = a\ (r,\ \theta,\ \varphi)\ e^{\frac{2\pi i}{h}Wt}$, one will have:

$$\Delta a + \frac{4\pi^2}{h^2 c^2}\left[E^2 + 2m_0 c^2 E + \frac{2e^2}{r}(m_0 c^2 + E)\frac{e^4}{r^2}\right]a = 0.$$

(21)

Upon posing the decomposition (3) here again, one will easily see that $Y\ (\theta,\ \varphi)$ must be a Laplace spherical function, and $R\ (r)$ must obey the equation:

$$\frac{d^2 R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[A + \frac{B}{r} + \frac{C}{r^2}\right]R = 0,$$

(22)

with the notations:

$$A = \frac{8\pi^2 m_0}{h^2}E\left(1 + \frac{E}{2m_0 c^2}\right),\ B = \frac{8\pi^2 m_0}{h^2}e^2\left(1 + \frac{E}{m_0 c^2}\right),\ C = -l\,(l+1) + \frac{4\pi^2 e^4}{h^2 c^2}.$$

(23)

The expression for C can also be written:

$$C = -l(l+1) + \alpha^2, \tag{23'}$$

in which $\alpha = \dfrac{2\pi e^2}{hc}$ is Sommerfeld's fine-structure constant.

We deduce from equation (22) that R has the asymptotic form $e^{-\rho/2}$, with $\rho = 2\sqrt{-A}\cdot r$, and upon setting:

$$R(\rho) = e^{-\frac{\rho}{2}}\, \upsilon(\rho), \tag{24}$$

if we substitute that in (22) then we will get:

$$\frac{d^2\upsilon}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{d\upsilon}{d\rho} + \left[\left(\frac{B}{2\sqrt{-A}} - 1\right)\frac{B}{\rho} + \frac{C}{\rho^2}\right]\upsilon = 0. \tag{25}$$

That is a linear differential equation that admits the point $\rho = 0$ as its only singularity at a finite distance. From the general theory of linear solutions, there exist, moreover, two solutions of (25) that are regular in a neighborhood of $\rho = 0$ and have the form:

$$\upsilon(\rho) = \rho^\gamma \sum_v a_v \rho^v \qquad (a_0 \neq 0). \tag{26}$$

The exponent γ (which is not necessarily an integer) is given by the defining equation:

$$\gamma(\gamma-1) + 2\gamma + C = \gamma(\gamma+1) + \alpha^2 - l(l+1) = 0, \tag{27}$$

which one can just as well write:

$$\left(\gamma + \frac{1}{2}\right)^2 = \left(l + \frac{1}{2}\right)^2 - \alpha^2. \tag{28}$$

One thus has the two values of γ:

$$\gamma = -\tfrac{1}{2} \pm \sqrt{\left(l + \tfrac{1}{2}\right)^2 - \alpha^2}\,. \tag{29}$$

One discards the solution that corresponds to the – sign by remarking that the corresponding function $\upsilon(\rho)$ will have a pole at $\rho = 0$. We then keep formula (29) with the + sign. It is appropriate to remark that even choosing the + sign will raise a small difficulty if $l = 0$, since that will then give a very small negative value for γ, and the function $\upsilon(\rho)$ will be infinite (granted, it will have a very small order) at $\rho = 0$. We agree to pass over that difficulty, which we shall encounter once more in Dirac's theory, because the function $\upsilon(\rho)$, although infinite at $\rho = 0$, is nonetheless square-summable.

If we introduce the form (26) into (25) then we will obtain the recurrence relation:

$$[(v + \gamma + 1)(v + \gamma) + 2\,(v + \gamma + 1) + C]\,a_{v+1} = \left[v + \gamma + 1 - \frac{B}{2\sqrt{-A}} \right] a_v .$$ (30)

The function R will be zero at infinity if all the a_v that are greater than a certain a_p are zero, which will happen if one has:

$$p + \gamma + 1 = \frac{B}{2\sqrt{-A}} .$$ (31)

Upon substituting the values of A, B, and γ into (31), one will find that:

$$\frac{\alpha\left(1 + \dfrac{E}{m_0 c^2}\right)}{\sqrt{-\dfrac{2E}{m_0 c^2}\left(1 + \dfrac{E}{2m_0 c^2}\right)}} = \left(p + \tfrac{1}{2}\right)^2 + \sqrt{\left(l + \tfrac{1}{2}\right)^2 - \alpha^2} ,$$ (32)

so

$$\frac{\alpha^2}{\left[\left(p + \tfrac{1}{2}\right)^2 + \sqrt{\left(l + \tfrac{1}{2}\right)^2 - \alpha^2} \right]^2} = \frac{-\dfrac{2E}{m_0 c^2}\left(1 + \dfrac{E}{2m_0 c^2}\right)}{\left(1 + \dfrac{E}{m_0 c^2}\right)^2} .$$ (33)

Upon adding one to both sides of this and taking inverses, one will easily get:

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{\left[\left(p + \tfrac{1}{2}\right)^2 + \sqrt{\left(l + \tfrac{1}{2}\right)^2 - \alpha^2} \right]^2} \right]^{-1/2} .$$ (34)

This is a formula that is quite analogous to the old Sommerfeld formula [formula (38) of Chapter I], but in which the whole numbers n_2 and n_1 have been replaced with:

$$p + \tfrac{1}{2} \quad \text{and} \quad l + \tfrac{1}{2},$$

respectively.

If we define a total quantum number n by the relation:

$$n = \left(p + \tfrac{1}{2}\right) + \left(l + \tfrac{1}{2}\right) = p + l + 1,$$ (35)

and develop (34) in powers of α^2, while neglecting terms of order higher than two, then we will obtain the approximate formula:

$$E_n = -\frac{Rh}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right)\right], \tag{36}$$

in which R = Rydberg constant. Formula (36) should be compared with formula (41) of Chapter I.

When one repeats the same calculation for an atom of rank N that has been ionized (N – 1) times, one will find:

$$E_{nl} = -\frac{RhN^2}{n^2}\left[1+\frac{\alpha^2 N^2}{n^2}\left(\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right)\right], \tag{37}$$

instead of (36), which is analogous to formula (44) in Chapter I.

In the case of X-rays, if one limits oneself to considering the inner electrons crudely as forming an electrostatic screen then one will get:

$$E_{nl} = -\frac{Rh(N-z_{nl})}{n^2}\left[1+\frac{\alpha^2(N-z_{nl})}{n^2}\left(\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right)\right], \tag{38}$$

and that formula will give the fine structure of the X-ray spectral terms, which should be compared with the old formula (9) in Chapter III.

Unfortunately, the agreement between the old Sommerfeld formulas and experiments is not recovered here. For example, take the doublets in the Balmer series. In Bohr's original model, a line in the Balmer series was generated during the transition of the H atom from an initial state of energy E_i to a final state that is characterized by $n = 2$. For Sommerfeld, the quantum number $n = 2$ corresponds to two possible values of the azimuthal quantum number $k = 1$ and $k = 2$. Hence, one has a fine structure for each line that is, in reality, a doublet, so the frequency width of the doublet must be, from formula (40) of Chapter I, equal to:

$$\delta v = \frac{1}{h}(E_{22}-E_{21}) = \frac{R\alpha^2}{2^4}\left(\frac{2}{1}-\frac{2}{2}\right) = \frac{R\alpha^2}{16} = 0.36 \text{ cm}^{-1}. \tag{39}$$

That number is in good agreement with experiment, if not somewhat better. With our new formula (36), we must attribute the quantum numbers $n = 2$, $l = 0$, and $n = 2$, $l = 1$ to the two levels that are responsible for the Balmer doublets, and we will find that:

$$\delta v = \frac{R\alpha^2}{2^4}\left(\frac{2}{1/2}-\frac{2}{3/2}\right) = \frac{R\alpha^2}{16}\cdot\frac{8}{3}. \tag{40}$$

We then find the 8 / 3 of the old number, which is already a bit too large!

Similarly, one will also find a number that is too large from formula (38) for the width of the Sommerfeld doublets for the X-rays.

Moreover, the difficulty that was pointed out before still remains here: It is certain (at least, for X-rays) that the levels that are responsible for Sommerfeld's fine structure have

the same quantum number k (i.e., the same number $l = k - 1$) and differ by a quantum number j that the preceding theory ignored completely.

Therefore, the relativistic wave mechanics that was developed in the preceding chapter is shown to be inadequate, since it introduces new difficulties without resolving the old ones.

3. Wave mechanics with one wave function and the Zeeman Effect. – We shall now see that the original wave mechanics (with one wave function) cannot account for the anomalous Zeeman Effect.

Consider an atom that is embedded in a uniform magnetic field **H**. Take the direction of the field to be the z-axis. We can then write the vector potential from which **H** is derived in the form:

$$A_x = -\frac{Hy}{2}, \qquad A_y = \frac{Hx}{2}, \qquad A_z = 0, \qquad (41)$$

since the relation **H** = rot **A** will then give $H_x = H_y = 0$ and $H_z = H$. The relativistic wave equation (17) of the last chapter can then be written ($\varepsilon = -e$):

$$\Delta\Psi - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \frac{4\pi i}{h}\frac{eV}{c^2}\frac{\partial\Psi}{\partial t} + \sum_{x,y,z}\frac{4\pi i}{h}\frac{e}{c}A_x\frac{\partial\Psi}{\partial x} = \frac{4\pi^2}{h^2}\left[m_0^2 c^2 - \frac{e^2}{c^2}(V^2 - \mathbf{A}^2)\right]\Psi = 0. \quad (42)$$

We shall prove the following theorem, which is nothing but the transposition of Larmor's classical theorem, which was recalled in Chapter IV, para. 2, into wave mechanics:

Theorem. – *When the field **H** is very weak, and one can neglect the relativistic corrections, the wave equation of the atom, when expressed in a reference that rotates around the direction of the field **H** with the angular velocity:*

$$o = \frac{1}{2}\frac{eH}{m_0 c}, \qquad (43)$$

*will be the same as if the reference system did not rotate and the field **H** did not exist.*

From our hypotheses, we suppose that **H** is small and $\mathbf{H}^2$ is negligible. Similarly, we suppose that $\eta = \dfrac{E}{m_0 c^2}$ is small (i.e., the Newtonian approximation) and η^2 is negligible. Finally, we also neglect the product ηH.

We can always take a direct system of axes $oxyz$. The oz axis will serve as the cylinder axis, and we define cylindrical coordinates z, ρ, φ by the usual formulas:

$$x = \rho\cos\varphi, \quad y = \rho\sin\varphi, \quad z = z. \qquad (44)$$

Thanks to (41), one will then have:

$$\frac{4\pi i}{h}\frac{e}{c}\sum_{x,y,z}A_x\frac{\partial\Psi}{\partial x}=\frac{2\pi i}{h}\frac{e}{c}H\left(y\frac{\partial\Psi}{\partial x}-x\frac{\partial\Psi}{\partial y}\right)=-\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi}, \qquad (45)$$

and equation (42) can be written in cylindrical coordinates as:

$$\frac{\partial^2\Psi}{\partial z^2}+\frac{\partial^2\Psi}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial\Psi}{\partial\rho}+\frac{1}{\rho^2}\frac{\partial^2\Psi}{\partial\varphi^2}-\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2}+\frac{4\pi i}{h}\frac{eV}{c^2}\frac{\partial\Psi}{\partial t}-\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi}$$
$$=\frac{4\pi^2}{h^2}\left(m_0^2-\frac{e^2V^2}{c^2}\right)\Psi, \qquad (46)$$

if one neglects $\mathbf{A}^2$.

Now take a system of cylindrical coordinates z', ρ', φ' that rotates around oz with angular velocity (43). One will have:

$$z'=z, \quad \rho'=r, \quad \varphi'=\varphi-ot, \quad t'=t, \qquad (47)$$

and as a result (upon neglecting o^2 in the order of $\mathbf{H}^2$):

$$\frac{\partial}{\partial z}=\frac{\partial}{\partial z'}, \qquad \frac{\partial}{\partial\rho}=\frac{\partial}{\partial\rho'}, \qquad \frac{\partial}{\partial t}=\frac{\partial}{\partial t'}-o\frac{\partial}{\partial\varphi}, \qquad \frac{\partial^2}{\partial t^2}=\frac{\partial^2}{\partial t'^2}-2o\frac{\partial^2}{\partial\varphi\,\partial t'}. \qquad (47')$$

Equation (46) can then be written:

$$\frac{\partial^2\Psi}{\partial z'^2}+\frac{\partial^2\Psi}{\partial\rho'^2}+\frac{1}{\rho'}\frac{\partial\Psi}{\partial\rho'}+\frac{1}{\rho'^2}\frac{\partial^2\Psi}{\partial\varphi'^2}-\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t'^2}+\frac{4\pi i}{h}\frac{eV}{c^2}\frac{\partial\Psi}{\partial t'}$$
$$+\frac{2o}{c^2}\frac{\partial^2\Psi}{\partial\varphi'\partial t'}-\frac{4\pi i}{h}o\frac{\partial\Psi}{\partial\varphi'}-\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi'}$$
$$=\frac{4\pi^2}{h^2}\left(m_0^2c^2-\frac{e^2V^2}{c^2}\right)\Psi \qquad (48)$$

in the system of primed variables.

Now, upon taking (43) into account and neglecting $o\eta$, one will have:

$$\frac{2o}{c^2}\frac{\partial^2\Psi}{\partial\varphi'\partial t'}=\frac{2o}{c^2}\frac{\partial}{\partial\varphi'}\left[\frac{2\pi i}{h}m_0c^2(1+\eta)\Psi\right]=\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi'}. \qquad (49)$$

The term $\dfrac{2o}{c^2}\dfrac{\partial^2\Psi}{\partial\varphi'\partial t'}$ then compensates for the last term in the left-hand side of (48), in the assumed approximation. As in the Newtonian approximation, the potential energy

must be regarded as being very small in comparison to the internal energy $m_0 c^2$, so the term $\dfrac{4\pi i}{h} \dfrac{eV}{c^2} \dfrac{\partial \Psi}{\partial \varphi'}$ will be negligible in comparison to the term $\dfrac{4\pi i}{h} \dfrac{eV}{c^2} \dfrac{\partial \Psi}{\partial t}$.

Finally, equation (48) reduces to:

$$\frac{\partial^2 \Psi}{\partial z'^2} + \frac{\partial^2 \Psi}{\partial \rho'^2} + \frac{1}{\rho'} \frac{\partial \Psi}{\partial \rho'} + \frac{1}{\rho'^2} \frac{\partial^2 \Psi}{\partial \varphi'^2} - \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t'^2} + \frac{4\pi i}{h} \frac{eV}{c^2} \frac{\partial \Psi}{\partial t'} = \frac{4\pi^2}{h^2} \left(m_0^2 c^2 - \frac{e^2 V^2}{c^2} \right) \Psi . \quad (50)$$

That equation is the same as if the primed system did not turn and the field **H** did not exist: Larmor's theorem is then proved.

We shall now recover the normal Zeeman Effect directly. Indeed, in the absence of an external magnetic field, the stationary state that is characterized by the two quantum numbers n and l, and an energy of E_{nl}^H will correspond to a wave function of the form:

$$\Psi (z, \rho, \varphi, t) = F (\rho, z) \, e^{im\varphi} \, e^{\frac{2\pi i}{h}(m_0 c^2 + E_{nl}^0)t} \quad (51)$$

in the fixed system z, ρ, φ, in which m is a positive or negative integer. By virtue of the theorem that was proved above, the function Ψ will have the same form that it would have in the system z', ρ', φ', in the presence of the field **H**. One will then have:

$$\Psi (z', \rho', \varphi', t') = F'(\rho', z') \, e^{im\varphi'} \, e^{\frac{2\pi i}{h}(m_0 c^2 + E_{nl}^0)t'} . \quad (52)$$

By virtue of (47), the wave function will have the expression:

$$\Psi (z, \rho, \varphi, t) = F (\rho, z) \, e^{im(\varphi - ot)} \, e^{\frac{2\pi i}{h}(m_0 c^2 + E_{nl}^0)t} = F (\rho, z) \, e^{im\varphi} \, e^{\frac{2\pi i}{h}(m_0 c^2 + E_{nl}^0 - \frac{mh}{2\pi}o)t} \quad (53)$$

in the fixed system.

The energy E_{nl}^H of the atom in the fixed system in the presence of the field **H** is then:

$$E_{nl}^H = E_{nl}^0 - m \frac{ho}{2\pi} = E_{nl}^0 - m \frac{eh H}{4\pi m_0 c} . \quad (54)$$

The spectral terms are then modified by the addition of a multiple of $\dfrac{eh H}{4\pi m_0 c}$, and the line that had the frequency:

$$v = \frac{1}{h} \left(E_{n'l'}^0 - E_{nl}^0 \right)$$

will now have the frequency:

$$v + (m - m') \frac{eh H}{4\pi m_0 c}$$

in the presence of the field. It will suffice to assume the selection rule $\delta m = 0, \pm 1$, which we shall justify later, in order to recover the normal Zeeman Effect, but there is no trace of the anomalies or the Landé factor in this theory. Wave mechanics with one wave function then gives nothing more than the classical theory or the old quantum theory.

4. "Selection rules" in wave mechanics. – In classical electromagnetism, when a distribution of electricity is animated with an accelerated motion, it will emit electromagnetic radiation. If $\rho\,(x,\,y,\,z,\,t)$ is the density of the electrical distribution then the intensity of the emitted radiation will be determined, in the first approximation, by the electric moment; i.e., by the vector **m** that is given by:

$$m_x = \iiint \rho x \, dx \, dy \, dz \,, \quad m_y = \iiint \rho y \, dx \, dy \, dz \,, \quad m_z = \iiint \rho z \, dx \, dy \, dz \,, \tag{55}$$

in rectangular components.

If one supposes that the quantities (55) are developed in a Fourier series in the form:

$$m_x = \text{cont.} + \sum_i m_x^i \cos(2\pi v_i t + \varepsilon_i) \tag{56}$$

then the energy that is emitted per second by radiation of frequency v_i whose electrical vibration is parallel to the x-axis will be equal to $\dfrac{64\pi^4}{3c^2} v_i^4 (m_x^i)^2$.

Knowing the electric moment of a distribution will then permit one to predict the frequencies, polarizations, and intensities of the emitted radiation.

According to the classical way of thinking, radiation is emitted in a continuous fashion, and radiated energy is progressively borrowed from the motion of the electric distribution, which will be a motion that dies away as a consequence. Things present themselves very differently in quantum theory: Radiation is emitted in the form of a quantum hv during the transition from one quantum state to the other. In order to evaluate the intensity of the energy that is emitted in the form of radiation by an assemblage of atomic emitters, one must argue statistically as follows: If N_n is the number of atoms in the energy state E_n then there will be a certain probability $P_{nm}\,dt$ that one of those atoms will pass to the state with energy E_m during the time interval dt, and the energy that is radiated per second will be (if one supposes that N_n is very large, so its decrease will be negligible):

$$N_n \, P_{nm} \, (E_n - E_m) = N_n \, P_n \, hv_{nm} \,. \tag{57}$$

The problem is to evaluate the probability P_{nm}. In order to do that, Bohr started with a very fertile idea, namely, correspondence. He assumed that one must recover the classical laws asymptotically for very close stationary states that correspond to large values of quantum numbers, and one can then state certain rules for the prediction of intensities and polarization. The development of the new mechanics has permitted us to make the rules that Bohr presented more precise.

In order to understand how one can arrive at the new rules for predicting the intensities and polarizations, we begin by arguing as Schrödinger first did. Consider an atomic system with one electron whose wave function is:

$$\Psi = \sum_n c_n\, a_n(x, y, z)\, e^{\frac{2\pi i}{h}E_n t}. \tag{58}$$

We have seen that one can associate it with the mean electricity distribution that is defined by the density $\delta = -\varepsilon\,\Psi\Psi^*$; the electric moment that corresponds to it is:

$$m_x = -e \iiint x\,\Psi\Psi^* d\tau = \text{const.} + \sum_{n,m}(-e)c_n\, c_m^* \iiint x\, a_n\, a_m^*\, d\tau\, e^{\frac{2\pi i}{h}(E_n - E_m)t}, \tag{59}$$

etc. Now, the integral that appears in the right-hand side of the last equation in (59) is the with indices mn of the matrix that is generated by the operator x in the system of a_i. If one denotes it by X_{mn} then one can write (59) in the form:

$$m_x = \text{const.} + 2\sum_{n,m}^{n \neq m}(-e)\cdot|c_n|\cdot|c_m|\cdot|X_{mn}|\cos 2\pi\left[\frac{E_n - E_m}{h}t + \varphi_m - \varphi_n\right]. \tag{60}$$

The frequencies that appear in the development (60) are precisely the Bohr frequencies $v_{mn} = (E_n - E_m)\,/\,h$. It is then entirely natural to think that a set of atoms that is found in the state that is defined by the wave function (58) will emit a line of frequency v_{mn} with an intensity that is proportional to $v_{mn}{}^2\,|X_{mn}|^2$.

However, as we saw, the fact that the radiation is emitted as quanta that are linked with the transitions between stationary states obliges us to depart from the classical image of an electric distribution of radiation of density $-\,e\,\Psi\Psi^*$ that radiates all of the Bohr frequencies simultaneously and continuously. Conforming to the spirit of the new mechanics itself, we are constrained to adopt a purely statistical statement. Here is the statement that we must adopt in order to avoid any contradiction:

Let there be an assemblage of identical atoms for which one has proper functions a_n for the Hamiltonian operator, and as a result, components of the matrix that take the form:

$$X_{mn} = \int a_m^*\, x \cdot a_n\, d\tau.$$

Suppose that there are N_n atoms in the state with energy E_n in that assemblage. The quantity of energy that is emitted per second by the set of atoms in the form of radiation of frequency $v_{mn} = (E_n - E_m)\,/\,h$ that vibrates electrically parallel to the x-axis will then be equal to $N_n\dfrac{64\pi^4}{3c^3}v_{nm}^4\,|X_{mn}|^2.$

Upon comparing this with (57), one will see that the probability $[P_{nm}]_x$ for the transition that produces the radiation in question will be given by:

$$[P_{nm}]_x = \frac{64\pi^4}{3c^3} \frac{v_{nm}^3}{h} \,|\, X_{mn} \,|^2 .\tag{61}$$

That is then the transition probability per unit time for an atom in the state E_n to go to the state E_m with emission of radiation whose vibration is parallel to the x-axis. When the three probabilities $[P_{nm}]_x$, $[P_{nm}]_y$, and $[P_{nm}]_z$ are found to be zero simultaneously for the two states E_n and E_m, the emission of radiation with frequency v_{mn} will not actually happen. One shows that the same thing will be true for all transitions that do not simultaneously satisfy the relations:

$$\delta l = 1, \qquad \delta m = \begin{cases} 0 \\ \pm 1 \end{cases}.\tag{62}$$

Wave mechanics then permits us to justify the selection rules for the quantum numbers l and m. Of course, it cannot justify the third selection rule that we encountered $\delta j = \begin{cases} 0 \\ \pm 1 \end{cases}$, since it ignores the number j. We shall see that Dirac's theory permits one to recover all three selection rules.

_____

PART II

—

GENERAL PRINCIPLES OF
DIRAC'S THEORY OF THE MAGNETIC
AND ROTATING ELECTRON

CHAPTER IX

PAULI'S ATTEMPT

1. Electron rotation and the polarization of light

We saw that wave mechanics, in its original form, did not resolve any of the problems that one was led to attach to the existence of proper magnetism for the electron. It has then become obvious that wave mechanics will remain incomplete as long as it does not contain an element that would correspond to that proper magnetism. However, the idea of a spinning, magnetic electron cannot be developed as easily in the new theory as it could in the old one. Remaining in the context of classical concepts, Uhlenbeck and Goudsmit imagined the classical electron to be a small sphere of negative electricity in rotation that presented a angular momentum that was equal to $\frac{1}{2} \cdot \frac{h}{2\pi}$ and a magnetic moment with the same axis that is equal to a Bohr magneton $\frac{eh}{4\pi m_0 c}$. However, in wave mechanics, it seems to be forbidden to make a representation that is also precise, and one must always introduce the language of probabilities.

In order to have some idea of the manner by which one must pose the problem, it is good to reflect on the manner by which one defines the polarization of the light quantum using the new ideas. Consider a sheaf of rectilinearly polarized light that propagates in the oz direction.

Let ox and oy be two axes that are perpendicular to oz. The normal vector to oz has the form:

$$\mathbf{a}_0 \sin 2\pi v \left(t - \frac{z}{c} + \varphi \right),$$

and it has components along ox and oy whose amplitudes are given by the formulas:

$$a_x = a_0 \cos \theta, \qquad a_y = a_0 \cos \theta, \tag{1}$$

if one lets θ denote the angle between that vector and ox; the light intensity of the sheaf is a_0^2.

If one places a Nicol prism that lets only vibrations that are parallel to ox pass in the path of the sheaf then the light vector after the Nicol will be reduced to $a_x \sin 2\pi v \left(t - \frac{z}{c} + \varphi \right)$, and the intensity will be $a_0^2 \cos^2 \theta$. If one then turns the Nicol by 90° then the light vector in the transmitted vibration will be parallel to oy and equal to:

$$a_y \sin 2\pi v \left(t - \frac{z}{c} + \varphi \right);$$

the transmitted intensity will be $a_0^2 \sin^2 \theta$.

How is one to interpret all of this if one is to assume the existence of photons? If one attributes a polarization to the photons individually then the causal description of the phenomenon will become impossible. Indeed, since one is, in some way, constrained to attribute the polarization that is defined by the light vector of the incident wave to the incident photon, one cannot perceive how it would be possible to explain causally why certain photons do not traverse the Nicol, while other ones do traverse it by taking a polarization that is directed along ox, for example. Upon adopting the probabilistic viewpoint of the new mechanics, one must consider that one cannot attribute a well-defined polarization to the incident photon, but one must only define the probability for the photon to reveal itself as having a polarization that is parallel to ox upon traversing the Nicol with the aid of the associated light wave; that probability will be $\cos^2 \theta$. One can say that before traversing the Nicol the photon had a probability of $\cos^2 \theta$ that it would reveal itself to be polarized parallel to ox and a probability of $\sin^2 \theta$ that it would reveal itself to be polarized parallel to oy.

Now return to the electron spin. The hypothesis of the spinning, magnetic electron leads one to associate two directed magnitudes with the electron that are proportional to each other, namely, the proper magnetic moment and the proper angular momentum. Whereas the polarization of the photon is defined by a direction that has no sense to it, the spin vector for the electron has both a direction and a sense. However, just as we cannot attribute a definite polarization to the photon, we cannot attribute a definite spin to an individual electron in a theory of the magnetic electron that conforms to the general principles of the new mechanics. The only thing that we can speak of is the probability that an experiment that permits us to determine the direction of the spin of an electron will give this or that result.

It was upon that fundamental idea that Pauli based a first attempt at a theory of the magnetic electron in the context of wave mechanics. We must elaborate a little bit upon that attempt, because it guided Dirac in the development of his more complete theory.

2. Pauli's theory. – Pauli ([19]) made a first attempt to neatly pose the problem of the magnetic electron in the context of the general ideas of the new mechanics.

If one considers a system of rectangular axes then any experiment that permits one to assign a value to the component of the moment of proper rotation of an electron along oz will give either $+\frac{1}{2} \cdot \frac{h}{2\pi}$ or $-\frac{1}{2} \cdot \frac{h}{2\pi}$ for a result, from the Uhlenbeck-Goudsmit picture. Moreover, it is natural to suppose that one must associate the electron, not with one function Ψ, but two functions $\Psi_1(x, y, z, t)$ and $\Psi_2(x, y, z, t)$, in such a way that $|\Psi_1|^2 \, dx \, dy \, dz$ measures the probability for the coordinates of the electron at time t to be found in the interval $dx \, dy \, dz$ and for the z component of its angular momentum to be $+\frac{1}{2} \cdot \frac{h}{2\pi}$, while $|\Psi_2|^2 \, dx \, dy \, dz$ measures the probability for the coordinates at time t to be found in

---

([19]) Zeitschrift für Physik **43** (1927), pp. 601.

the interval $dx\,dy\,dz$ and the z component of the angular momentum to be $-\dfrac{1}{2}\cdot\dfrac{h}{2\pi}$. Pauli therefore introduced the idea that in order to take into account the magnetism of the electron, one must increase the number of functions Ψ. Naturally, the normalization condition will become:

$$\iiint \left[\,|\Psi_1|^2 + |\Psi_2|^2\,\right]\,dx\,dy\,dz = 1 \tag{2}$$

here.

Pauli looked for the equation of propagation that would determine the two wave functions Ψ. In order to do that, he always started with the Hamiltonian function, but while taking the magnetic moment of the electron into account.

If one denotes the components of a unit vector **s** that points in the direction of the electron spin by s_x, s_y, s_z then the magnetic moment will be $\dfrac{eh}{4\pi m_0 c}\cdot\mathbf{s}$.

The Hamiltonian function in a constant field will then be of the form $H(x,\,y,\,z,\,p_x,\,p_y,\,p_z,\,s_x,\,s_y,\,s_z)$, and the quantities s_x, s_y, s_z will appear linearly in it, as is easy to see. Now, the artifice that permits wave mechanics to obtain the equation of propagation consists of replacing p_x, p_y, p_z with the operators $-\dfrac{h}{2\pi i}\cdot\dfrac{\partial}{\partial x}$, etc., and writing:

$$H\left(x,\,y,\,z,\,-\frac{h}{2\pi i}\frac{\partial}{\partial x},\,-\frac{h}{2\pi i}\frac{\partial}{\partial y},\,-\frac{h}{2\pi i}\frac{\partial}{\partial z},\,s_x,\,s_y,\,s_z\right)\Psi = E\Psi \tag{3}$$

for the monochromatic waves of frequency $E\,/\,h$. However, we must have two equations for the two Ψ. Guided by some general considerations that I will tacitly pass over, Pauli imagined replacing the components s_x, s_y, and s_z, respectively, with three Hermitian matrices with 2 rows and 2 columns:

$$\mathbf{s}_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix},\qquad \mathbf{s}_2 = \begin{vmatrix} 0 & -i \\ +i & 0 \end{vmatrix},\qquad \mathbf{s}_3 = \begin{vmatrix} +1 & 0 \\ 0 & -1 \end{vmatrix}. \tag{4}$$

Moreover, he agreed to pose the following definition: If A is a matrix with 2 rows and 2 columns then the operation $A\Psi$ will be defined by the relation:

$$A\,\Psi_i = \sum_{k=1}^{2} a_{ik}\Psi_k \qquad (i,\,k = 1,\,2), \tag{5}$$

which is a natural generalization of some formulas that we encountered before that were concerned with matrices. In particular, an application of the relation (5) to the matrices $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$ will give:

$$\begin{aligned} \mathbf{s}_1\Psi_1 &= \ \ \ \Psi_2\,, & \mathbf{s}_1\Psi_2 &= \ \ \ \Psi_1\,, \\ \mathbf{s}_2\Psi_1 &= -\,i\Psi_2\,, & \mathbf{s}_2\Psi_2 &= \ \ i\Psi_1\,, \\ \mathbf{s}_3\Psi_1 &= \ \ \ \Psi_1\,, & \mathbf{s}_3\Psi_2 &= -\,\Psi_2\,. \end{aligned} \tag{6}$$

Pauli then replaced equation (3) with the two equations:

$$H\left(x,y,z,-\frac{h}{2\pi i}\frac{\partial}{\partial x},-\frac{h}{2\pi i}\frac{\partial}{\partial y},-\frac{h}{2\pi i}\frac{\partial}{\partial z},\mathbf{s}_x,\mathbf{s}_y,\mathbf{s}_z\right)\Psi_1 = E\Psi_1,$$

(7)

$$H\left(x,y,z,-\frac{h}{2\pi i}\frac{\partial}{\partial x},-\frac{h}{2\pi i}\frac{\partial}{\partial y},-\frac{h}{2\pi i}\frac{\partial}{\partial z},\mathbf{s}_x,\mathbf{s}_y,\mathbf{s}_z\right)\Psi_2 = E\Psi_2.$$

These are two simultaneous equations that the functions Ψ_1 and Ψ_2 must satisfy.

However, the theory that was developed above makes the oz axis play a special role, since the wave functions Ψ_1 and Ψ_2 give the probabilities of the two possible values for the component of the spin that is *parallel* to oz. If one wishes to obtain the probabilities for the component of the angular momentum that is parallel to another arbitrary axis OZ to be equal to $+\frac{1}{2}\cdot\frac{h}{2\pi}$ or $-\frac{1}{2}\cdot\frac{h}{2\pi}$ then one must take another system of rectangular axes $OXYZ$ in which OZ is the third axis and then calculate the two wave functions in that system – say, $\Phi_1(X, Y, Z)$ and $\Phi_2(X, Y, Z)$ – whose moduli squared will give the desired probabilities, and naturally, since nothing will physically distinguish the system of axes $OXYZ$ from system $oxyz$, the functions Φ must obey the following equations, which one will obtain immediately from (7) by transposition:

$$H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{s}_x,\mathbf{s}_y,\mathbf{s}_z\right)\Phi_1 = E\Phi_1,$$

(8)

$$H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{s}_x,\mathbf{s}_y,\mathbf{s}_z\right)\Phi_2 = E\Phi_2.$$

The position of the axes $OXYZ$ with respect to the axes $oxyz$ can be defined in the following fashion: If we replace x, y, z, X, Y, Z with x_1, x_2, x_3, X_1, X_2, X_3, respectively, then we will have transformation formulas of the form:

$$X_i = \sum_j o_{ij} x_j,$$

(9)

in which the o_{ij} are elements of a real, orthogonal matrix $\left(\sum_j o_{ij} o_{ik} = \delta_{jk}\right)$. In short, the matrix O defines the passage from the first system of axes to the second one.

If that matrix is known then how can one express the functions Φ with the aid of the functions Ψ? That was the question that Pauli posed and solved. If one starts with equations (7) and one replaces x, y, z as functions of X, Y, Z with formulas that are inverse to (9) then, as we will verify later on in a special case, one will obtain the equations:

$$H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{S}_1,\mathbf{S}_2,\mathbf{S}_3 \right)\Psi_1(X,\,Y,\,Z) = E\,\Psi_1(X,\,Y,\,Z),$$

$$\tag{10}$$

$$H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{S}_1,\mathbf{S}_2,\mathbf{S}_3 \right)\Psi_2(X,\,Y,\,Z) = E\,\Psi_2(X,\,Y,\,Z).$$

The operator H in equations (10) is defined in the following manner: One writes the classical Hamiltonian function:

$$H(X,\,Y,\,Z,\,p_X,\,p_Y,\,p_Z,\,s_X,\,s_Y,\,s_Z)$$

in the system XYZ, which is a function in which appear the components s_X, s_Y, and s_Z of the unit vector that points along the proper magnetic moment of the electron. One then replaces, on the one hand, p_X, p_Y, p_Z with, $-\dfrac{h}{2\pi i}\dfrac{\partial}{\partial X}$, etc., resp., and on the other hand, s_X, s_Y, s_Z with the Hermitian matrices $\mathbf{S}_X$, $\mathbf{S}_Y$, $\mathbf{S}_Z$, resp., which are defined by the relations:

$$\mathbf{S}_i = \sum_j o_{ij}\mathbf{s}_j , \tag{11}$$

by starting with the $\mathbf{s}_i$ that are defined by (4). The $\mathbf{S}_i$ will always figure linearly in the operator H that is formed in this way. Finally, the functions $\Psi_1(X,\,Y,\,Z)$ and $\Psi_2(X,\,Y,\,Z)$ in formulas (10) are obtained by starting with $\Psi_1(x,\,y,\,z)$ and $\Psi_2(x,\,y,\,z)$ and replacing x, y, z with their expressions in terms of X, Y, Z.

Pauli then proved [and we shall carry out an analogous proof in Dirac's theory [20]] that that there always exists a unitary matrix Λ with two rows and two columns such that one will have:

$$\mathbf{S}_i = \Lambda^{-1}\,\mathbf{s}_i\,\Lambda, \qquad (\Lambda^{-1} = \Lambda^{+}) \qquad i = 1,\,2,\,3. \tag{12}$$

In other words, there exists a canonical transformation that takes each of the $\mathbf{s}_i$ to the corresponding $\mathbf{S}_i$. The matrix Λ corresponds to an operation that is defined by the formula (5), and upon applying that operation to the two sides of the equations (10), one will obtain:

$$\Lambda H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{S}_1,\mathbf{S}_2,\mathbf{S}_3 \right)\Psi_1 = E\,\Lambda\Psi_1,$$

$$\tag{13}$$

$$\Lambda H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{S}_1,\mathbf{S}_2,\mathbf{S}_3 \right)\Psi_2 = E\,\Lambda\Psi_2,$$

---

[20] Cf., Chapter XI, paragraph 2.

in which the Ψ are expressed in terms of X, Y, Z. Since the $\mathbf{S}_i$ figure linearly in H, ΛH contains the terms $\Lambda \mathbf{S}_i$ linearly, which are equal to $\mathbf{s}_i \Lambda$, from (12). Since Λ commutes with X, $\partial / \partial X$, etc., moreover, one will have:

$$\Lambda H(\ldots) \Psi_1 = H\left(X, Y, Z, -\frac{h}{2\pi i}\frac{\partial}{\partial X}, -\frac{h}{2\pi i}\frac{\partial}{\partial Y}, -\frac{h}{2\pi i}\frac{\partial}{\partial Z}, \mathbf{s}_x, \mathbf{s}_y, \mathbf{s}_z \right) \Psi_1 = E \Lambda \Psi_1, \quad (14)$$

and an analogous equation in Ψ_2. Upon comparing this with (8), one will see that one has:

$$\Phi_1(X, Y, Z) = \Lambda \Psi_1(X, Y, Z), \qquad \Phi_2(X, Y, Z) = \Lambda \Psi_2(X, Y, Z). \qquad (15)$$

Formulas (15) teach us how the wave functions Φ are deduced from the wave function Ψ because in each particular case if the matrix O is known then one can find the matrix Λ. We shall give a concrete example in the following paragraph.

Of course, since Λ is unitary, the sum $|\Phi_1|^2 + |\Phi_2|^2$ will be equal to $|\Psi_1|^2 + |\Psi_2|^2$, which expresses the fact that the total probability of the two possible hypotheses on the sense of the angular momentum will always be unity.

3. Example of an application of the preceding theory. – In order to illustrate the theory of the preceding paragraph, it is useful to study a very simple example that was developed by Pauli himself, namely, that of an electron at rest in a magnetic field $\mathbf{H}$.

First, take a system of axes $oxyz$ such that the positive direction of oz coincides with the direction of $\mathbf{H}$. One will then have $H_x = H_y = 0$, $H_z = H$. Upon introducing the Bohr magneton $\mu_0 = \dfrac{eh}{4\pi m_0 c}$, the Hamiltonian function will reduce to $\mu_0 s_z H$ [21], and one will have:

$$\mu_0 \mathbf{s}_3 H \Psi_1 = E \Psi_1, \qquad \mu_0 \mathbf{s}_3 H \Psi_2 = E \Psi_2 \qquad (16)$$

for equations (7), or, due to (6):

$$\mu_0 H \Psi_1 = E \Psi_1, \qquad -\mu_0 H \Psi_2 = E \Psi_2. \qquad (16)$$

This system will have a solution for only $E = \pm \mu_0 H$. Since $|\Psi_1|^2 + |\Psi_2|^2 = 1$, one will have:

$$\text{for } E = \mu_0 H: \qquad \Psi_1 = e^{i\gamma}, \qquad \Psi_1 = 0,$$
$$\text{for } E = -\mu_0 H: \qquad \Psi_1 = 0, \qquad \Psi_1 = e^{i\gamma}. \qquad (17)$$

Therefore, for $E = \mu_0 H$, the magnetic axis of the electron certainly points to the positive z and for $E = -\mu_0 H$, it will certainly point to the negative z. These are, in fact, the results that one must expect.

---

[21] In this paragraph, the letter H will no longer denote the Hamiltonian, but the magnitude of the magnetic field.

Now take a second system of axes *OXYZ* such that the *OZ* axis makes an angle θ with the field **H**. The classical Hamiltonian function will then be:

$$\mu_0 (\mathbf{s} \cdot \mathbf{H}) = \mu_0 (s_X H_X + s_Y H_Y + s_Z H_Z),$$

and, from (8), the wave functions Φ_1 and Φ_2 will be solutions of the equations:

$$\mu_0 [H_X \mathbf{s}_1 + H_Y \mathbf{s}_2 + H_Z \mathbf{s}_3] \Phi_1 = E \Phi_1,$$

$$\mu_0 [H_X \mathbf{s}_1 + H_Y \mathbf{s}_2 + H_Z \mathbf{s}_3] \Phi_2 = E \Phi_2,$$

(18)

which will give, upon making them explicit:

$$\mu_0 [(H_X - iH_Y) \Phi_2 + H_Z \Phi_1] = E \Phi_1,$$

$$\mu_0 [(H_X + iH_Y) \Phi_1 + H_Z \Phi_2] = E \Phi_2 .$$

(19)

In order for these two homogeneous equations in Φ_i to be compatible, it is necessary that:

$$\begin{vmatrix} \mu_0 H_X - E & \mu_0 (H_X - iH_Y) \\ \mu_0 (H_X + iH_Y) & -\mu_0 H_X - E \end{vmatrix} = 0;$$

(20)

i.e.:

$$E = \pm \mu_0 \sqrt{H_X^2 + H_Y^2 + H_Z^2} = \pm \mu_0 H .$$

(21)

Here, there are once more two cases to distinguish:

First case: $E = \mu_0 H$.

One will then have:

$$\frac{\Phi_2}{\Phi_1} = \frac{H - H_Z}{H_X - iH_Y} = \frac{H - H_Z}{H_X^2 + H_Y^2} (H_X + i H_Y) = \frac{H - H_Z}{H_P^2} \cdot H_P e^{i\alpha},$$

(22)

upon setting α = arc tan H_X / H_X and letting H_P denote the component of **H** that is normal to *OZ*. One will then also have:

$$\frac{\Phi_2}{\Phi_1} = \frac{H(1 - \cos \theta)}{H \sin \theta} e^{i\alpha} = \tan \frac{\theta}{2} \cdot e^{i\alpha}.$$

(23)

Since $| \Phi_1 |^2 + | \Phi_2 |^2 = 1$, one must have:

$$| \Phi_1 |^2 = \cos \frac{\theta}{2} e^{-i\frac{\alpha - \pi/2}{2}}, \qquad | \Phi_2 |^2 = i \sin \frac{\theta}{2} e^{i\frac{\alpha - \pi/2}{2}} .$$

(24)

Second case: $E = -\mu_0 H$.

One then has:

$$\frac{\Phi_2}{\Phi_1} = -\frac{H + H_Z}{H_X - iH_Y} = -\frac{H + H_Z}{H_P^2} \cdot e^{i\alpha} = -\frac{1 + \cos\theta}{\sin\theta} e^{i\alpha} = -\cot\frac{\theta}{2} \cdot e^{i\alpha}, \qquad (26)$$

and, in turn:

$$|\Phi_1|^2 = \frac{1}{1 + \cot^2\dfrac{\theta}{2}} = \sin^2\frac{\theta}{2}, \qquad |\Phi_2|^2 = \cos^2\frac{\theta}{2}, \qquad (27)$$

which permits one to set:

$$\Phi_1 = i\sin\frac{\theta}{2} \cdot e^{-i\frac{\alpha - \pi/2}{2}}, \qquad \Phi_2 = \cos\frac{\theta}{2} \cdot e^{i\frac{\alpha - \pi/2}{2}}. \qquad (28)$$

The results thus-obtained signify − to take an example − that if a first experiment shows us that the electron points in the direction of the field with the energy of $\mu_0 H$ then the probability that a second experiment will show us that it points in the *OZ* direction will be $\cos^2\theta/2$.

It is easy to calculate Pauli's unitary matrices Λ in the simple example that was treated here. Compare the functions Ψ in (17) (upon taking the arbitrary arguments γ and δ to be zero in order to the simplify things) with the functions Φ in (25) and (28). We see that the unitary matrix Λ is equal to:

$$\Lambda = \begin{vmatrix} \cos\dfrac{\theta}{2} \cdot e^{-\frac{\alpha - \pi/2}{2}} & i\sin\dfrac{\theta}{2} \cdot e^{-\frac{\alpha - \pi/2}{2}} \\ i\sin\dfrac{\theta}{2} \cdot e^{\frac{\alpha - \pi/2}{2}} & \cos\dfrac{\theta}{2} \cdot e^{\frac{\alpha - \pi/2}{2}} \end{vmatrix}, \qquad (29)$$

The elements Λ_{ij}, which are special cases of the Cayley-Klein parameters, satisfy the relation:

$$\Lambda_{11}\Lambda_{22} - \Lambda_{12}\Lambda_{21} = 1. \qquad (30)$$

One easily finds that:

$$\Lambda^+ = \begin{vmatrix} \Lambda_{22} & -\Lambda_{12} \\ -\Lambda_{21} & \Lambda_{11} \end{vmatrix}, \qquad (31)$$

from which, one infers that:

$$\Lambda\Lambda^+ = \begin{vmatrix} \Lambda_{22}\Lambda_{11} - \Lambda_{12}\Lambda_{21} & -\Lambda_{11}\Lambda_{12} + \Lambda_{12}\Lambda_{11} \\ \Lambda_{21}\Lambda_{22} - \Lambda_{22}\Lambda_{21} & \Lambda_{22}\Lambda_{11} - \Lambda_{12}\Lambda_{21} \end{vmatrix} = 1; \qquad (32)$$

Λ is thus indeed unitary. The form of Λ can be predicted by Pauli's argument, moreover.

4. Insufficiency of Pauli's theory. – We have given only a sketch of Pauli's theory. That theory is not entirely satisfactory, in fact. First of all, it is not consistent with the principle of relativity; viz., it only envisions spatial coordinate changes and not space-time coordinate changes in the relativistic sense. Moreover, it does not lead to a prediction of the hydrogen spectrum that is completely correct.

For these reasons, we shall not study it any further, but we must remark that it does introduce the following essential ideas:

1. The magnetism of the electron corresponds to the existence of several functions Ψ.

2. The wave functions Ψ must permit one to define the probability of the possible orientations for the spin in a certain direction.

3. One can preserve the form of the equations of propagation while changing the coordinate axes, but the wave functions will then transform in a certain manner that is defined by a matrix Λ.

Darwin ([22]) made another attempt to introduce the magnetism of the electron into wave mechanics in a manner that conformed to the principle of relativity: He sought to define four functions Ψ that were the components of a space-time vector. The attempt was not entirely successful, and he then lent his support to Dirac's theory, in which four functions Ψ also figure, but they are not the components of a space-time vector.

_____

([22]) Proceedings of the Royal Society A **116** (1927), pp. 227.

DIRAC'S THEORY

────

1. Review of previous results

In order to introduce the idea of the proper magnetism and proper rotation of the electron into the equations wave mechanics, Dirac started with some considerations that were more general than those of Pauli. He assumed that the equations of wave mechanics must be made to agree with the principle of special relativity, but he criticized the manner by which that agreement had been sought. In order to understand his criticisms, we shall first rapidly recall how we have written the equations of wave mechanics for a wave function Ψ.

The general equation of propagation for a corpuscle in non-relativistic wave mechanics is:

$$H\left(x, y, z, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}, t\right)\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t}, \tag{1}$$

in which H is the Hamiltonian operator. In the case of a corpuscle of mass m that displaces in a field that is characterized by a potential $U(x, y, z, t)$, one will have:

$$-\frac{h^2}{8\pi^2 m}\Delta\Psi + U(x, y, z, t)\,\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \tag{2}$$

If the corpuscle carries a charge $-\varepsilon$ and moves in an electrostatic field that is derivable from an electric potential $V(x, y, z, t)$ then one will have $U = \varepsilon V$, and (2) can be written:

$$-\frac{h^2}{8\pi^2 m}\Delta\Psi + \varepsilon V(x, y, z, t)\,\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t}. \tag{3}$$

In particular, if the charged corpuscle is an electron then one must set $\varepsilon = -e$ in (3).

We have seen that in the non-relativistic theory, the necessarily-positive quantity $\Psi\Psi^*$ represents the probability density of presence, and the total probability thus-defined is conserved (viz., equal to unity) in the course of time. The quantity $\varepsilon\Psi\Psi^*$ (which will be $-e\,\Psi\Psi^*$ for the electron) is the mean electric charge density that one uses to calculate the emitted radiation.

If one would like to construct a theory of wave mechanics that is accord with the principles of special relativity then one will immediately obtain a new equation of propagation that will replace (1) by a very natural induction. We have seen that for a corpuscle of mass m and charge ε that displaces in an electromagnetic field that is defined by a scalar potential $V(x, y, z, t)$ and a vector potential $\mathbf{A}(x, y, z, t)$, that new equation can be written:

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t}-\varepsilon V\right)^2\Psi-\sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}+\varepsilon A_x\right)^2\Psi=m_0^2c^2\Psi\ . \tag{4}$$

For an electron, equation (4) will take the form:

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t}+e V\right)^2\Psi-\sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}-\frac{e}{c} A_x\right)^2\Psi=m_0^2c^2\Psi\ . \tag{5}$$

We introduce the notations:

$$P_1=-\frac{h}{2\pi i}\frac{\partial}{\partial x}+\frac{e}{c}A_x\ ,\qquad\qquad P_2=-\frac{h}{2\pi i}\frac{\partial}{\partial y}+\frac{e}{c}A_y\ ,$$

$$\tag{6}$$

$$P_3=-\frac{h}{2\pi i}\frac{\partial}{\partial z}+\frac{e}{c}A_z\ ,\qquad\qquad P_4=\frac{h}{2\pi i}\frac{\partial}{\partial t}+\frac{e}{c}V\ .$$

That will permit us to replace equation (5) with:

$$\left(P_4{}^2-\sum_{i=1}^{3}P_i{}^2-m_0^2c^2\right)\Psi=0. \tag{7}$$

As we have seen, the relativistic equations of propagation (4)-(7) no longer permitted us to take the quantity $\Psi\Psi^*$ to be the probability density of localization because the integral of that quantity over all space was no longer necessarily constant in the course of time. We had to adopt a more complicated expression that was written:

$$\rho=-\frac{h}{4\pi m_0 c}\left(\Psi\frac{\partial\Psi^*}{\partial t}-\Psi^*\frac{\partial\Psi}{\partial t}\right)+\frac{e}{m_0 c^2}V\,\Psi\Psi^* \tag{8}$$

in the case of the electron ($\varepsilon=-e$), which is an expression that will reduce to $\Psi\Psi^*$ when the Newtonian approximation is sufficient. Here, the mean electric charge density will be $\rho\,\varepsilon=-\rho\,e$, with the value (8) for ρ.

2. Criticisms of equations (4)-(8) that were addressed by Dirac. – Dirac addressed some serious criticisms of the attempt that we just recalled at defining relativistic wave mechanics. In particular, he attacked the expression (8) for the probability density. Since that expression is imposed by the form itself of the equation of propagation (7), that would amount to attacking that equation itself.

We shall summarize Dirac's arguments. A first criticism that one can address in the expression (8) for ρ is that it is not necessarily positive-definite, although a negative value of ρ can obviously have no physical meaning. Another criticism that was pointed out already at the end of Chapter VII is the following one: The general principles of the

new mechanics demand that no matter what form the adopted equation of propagation might take, the probability that one finds for the coordinates of the corpuscle – i.e., of the values that are found in the intervals $x \to x + dx$, $y \to y + dy$, $z \to z + dz$ – will be $\Psi\Psi^* \, dx \, dy \, dz$, and that demand is not consistent with formula (8).

Faced with these difficulties, Dirac maintained that the probability density of localization must necessarily have the always-positive form $\Psi\Psi^*$, or if one (with Pauli) assumes the existence of several wave functions Ψ_i then it must necessarily have the likewise-always-positive form $\sum_i \Psi_i \Psi_i^*$. However, upon adopting that postulate, one will be led to a conclusion of paramount importance: The true equation (or equations) of propagation for relativistic wave mechanics must be of first-order with respect to the four variables x, y, z, t. We shall present his reasoning:

The non-relativistic equation (1) is of first order in time, and the corresponding expression for the probability density is $\Psi\Psi^*$. If one is given the initial form Ψ (x, y, z, 0) then it will be found to be determined completely by the equation of propagation, and as a result, one can say that when one gives the initial $\Psi\Psi^*$, the ultimate evolution of the density will be determined. One understands that the conservation of probability (and electricity) will then be a necessary consequence of the equation of propagation itself, as we have proved. Things are no longer the same in the case of the relativistic equation (4). Since that equation is of second order in t, one must, in fact, give values to Ψ and $\partial\Psi / \partial t$ at the initial instant in order for the wave function to be determined.

One must then adopt the expression (8) for the probability density if one wishes that conservation of probability should result from the equation of propagation. Now, adopt the postulate that the probability density must necessarily have the form $\Psi\Psi^*$ or $\sum_i \Psi_i \Psi_i^*$, as Dirac did. We shall see that the equation (or equations) of propagation must then be of first order in t, and that equation (4) cannot, in turn, be exact.

Indeed, if the equation of propagation is of second order in t then its solution will be determined only if one gives the initial values in Ψ and $\partial\Psi / \partial t$.

Suppose that one is given only the initial Ψ, but not the initial $\partial\Psi / \partial t$. With Dirac's hypothesis, the initial density will then be known, since it depends upon only Ψ, but the ultimate evolution of the function Ψ will not be. Upon giving an arbitrary initial form to $\partial\Psi / \partial t$, if one starts with a known initial form for the density then one can arrive at no particular ultimate state of evolution for that density, and the conservation of probability will not be automatic. Dirac then concluded that the equation of propagation would be of first order in t, and since the principle of relativity always makes the spatial coordinates and time play a symmetric role, it must be of first order with respect to the four variables, x, y, z, t. The same reasoning and the same conclusion are valid when there are several functions Ψ and several simultaneous equations of propagation.

Dirac was then compelled to seek one or more equations that would be of first order in the variables x, y, z, t in order to replace equation (4). We shall see how he succeeded in that quest.

3. The Dirac equations in the absence of a field. – In order to establish the new equations, Dirac started with the case of the free motion of the electron in the absence of any electromagnetic field.

One will then have $V = \mathbf{A} = 0$, and the operators P_i [formula (6)] will become:

$$p_1 = -\frac{h}{2\pi i}\frac{\partial}{\partial x}, \quad p_2 = -\frac{h}{2\pi i}\frac{\partial}{\partial y}, \quad p_3 = -\frac{h}{2\pi i}\frac{\partial}{\partial z}, \quad p_4 = -\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}.$$

In this simple case, one can assume with some certainty that the function Ψ will satisfy the equation:

$$\Delta\Psi - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} = \frac{4\pi^2}{h^2}m_0^2 c^2\Psi, \qquad \text{or}\ \left(p_4^2 - \sum_{i=1}^{3} p_i^2 - m_0^2 c^2\right)\Psi = 0, \qquad (10)$$

because in that case, it would result from fundamental considerations that grow out of wave mechanics that the monochromatic plane wave with frequency $\dfrac{W}{h} = \dfrac{m_0 c^2}{h\sqrt{1-\beta^2}}$ and

wavelength $\lambda = \dfrac{h}{p} = \dfrac{h\sqrt{1-\beta^2}}{m_0 v}$ would have to be a solution, and one would effortlessly

verify that the same thing would be true for equation (10).

However, equation (10) is of second order, and we would like to have equations of first order. In order to arrive at that, Dirac supposed that there existed several wave functions Ψ_1, Ψ_2, ..., Ψ_N and that each of the them satisfied the equation:

$$\left(p_4^2 - \sum_{i=1}^{3} p_i^2 - m_0^2 c^2\right)\Psi_k = 0 \qquad (k = 1, ..., N), \qquad (11)$$

but these equations of second order must be *consequences* of the true equations of propagation, which are of first order.

Dirac wrote those true first-order equations of propagation in the symbolic form:

$$(p_4 + \alpha_1\, p_1 + \alpha_2\, p_2 + \alpha_3\, p_3 + \alpha_4\, m_0\, c)\, \Psi = 0. \qquad (12)$$

The symbolic equation (12) signifies that one has:

$$\left(p_4 + \sum_{i=1}^{3}\alpha_i p_i + \alpha_4\, m_0\, c\right)\Psi_k = 0 \qquad (13)$$

for each Ψ_k.

The α_i are matrices with N rows and N columns, and one defines the operation $\alpha_i\,\Psi_k$, as we did already in our study of Pauli's theory, by the formula:

$$\alpha_i \, \Psi_k = \sum_{l=1}^{N} \alpha_{i,kl} \Psi_l \, , \tag{14}$$

in which $\alpha_{i,kl}$ denote the indices kl of the matrix α_i.

However, equations (13) must have equations (11) as a consequence, and that will impose certain conditions upon the matrices α_i. Indeed, apply the operator $p_4 - \sum_{i=1}^{3} \alpha_i p_i - \alpha_4 \, m_0 \, c$ to equation (13); we will find that:

$$\left[p_4^2 - \left(\sum_{i=1}^{3} \alpha_i p_i + \alpha_4 m_0 c \right)^2 \right] \Psi = 0, \tag{15}$$

and equation (15) will coincide with (11) only if one sets:

$$\alpha_i^{\ 2} = 1, \qquad \alpha_i \, \alpha_j + \alpha_j \, \alpha_i = 0. \tag{16}$$

We must then impose the conditions (16) on the matrices α_i. Moreover, we also impose the condition upon them that they must be Hermitian, like all of the matrices that enter into the new mechanics.

Dirac sought to make the number N of wave functions as small as possible. For $N < 4$, one cannot find four Hermitian matrices that satisfy the conditions (16). On the contrary, it is possible to find them for $N = 4$, and it will be convenient to adopt the following α_i matrices:

$$\alpha_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \ \alpha_2 = \begin{bmatrix} 0 & 0 & 0 & +i \\ 0 & 0 & -i & 0 \\ 0 & +i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix}, \ \alpha_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \ \alpha_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{17}$$

They are obviously Hermitian, and furthermore, they have the matrix 1 for their squares and anti-commute with each other, as the conditions (16) would desire.

Briefly, we (with Dirac) shall assume the existence of four functions Ψ_i that satisfy the four simultaneous first-order equations (13). These four equations can be written explicitly with the choice (17) of matrices α_i:

$$(p_4 + m_0 c) \, \Psi_1 + (p_1 + i \, p_2) \, \Psi_4 + p_3 \, \Psi_3 = 0,$$

$$(p_4 + m_0 c) \, \Psi_2 + (p_1 - i \, p_2) \, \Psi_3 - p_3 \, \Psi_4 = 0,$$

$$\tag{18}$$

$$(p_4 - m_0 c) \, \Psi_3 + (p_1 + i \, p_2) \, \Psi_2 + p_3 \, \Psi_1 = 0,$$

$$(p_4 - m_0 c) \, \Psi_4 + (p_1 - i \, p_2) \, \Psi_1 - p_3 \, \Psi_1 = 0.$$

Upon starting with the system (18), one can see in detail how that system has the second-order system (11) (with $N = 4$) as a consequence.

For example, apply the operator $p_4 - m_0 c$ to the first equation (18), and the operator $p_1 + i\, p_2$, which commutes with the preceding one, to the fourth equation; one will obtain:

$$(p_4^2 - m_0^2 c^2)\, \Psi_1 + (p_4 - m_0 c)\, (p_1 + i\, p_2)\, \Psi_4 + (p_4 - m_0 c)\, p_3\, \Psi_3 = 0,$$

$$\hspace{10cm} (19)$$

$$(p_1 + i\, p_2)\, (p_4 - m_0 c)\, \Psi_4 + (p_1^2 + p_2^2)\, \Psi_1 - (p_1 + i\, p_2)\, p_3\, \Psi_2 = 0,$$

so the result of eliminating Ψ_4 will be:

$$(p_4^2 - m_0^2 c^2)\, \Psi_1 + (p_1^2 + p_2^2)\, \Psi_1 + (p_1 + i\, p_2)\, p_3\, \Psi_2 + (p_4 - m_0 c)\, p_3\, \Psi_3 = 0. \qquad (20)$$

However, upon applying the operator p_3 to the third equation (18), one will find:

$$p_3\, (p_4 - m_0 c)\, \Psi_3 = -\, p_3\, (p_1 + i\, p_2)\, \Psi_2 - p_3^2\, \Psi_1, \qquad (21)$$

and since the p_i commute with each other, upon substituting this in (20), one will obtain the equation:

$$\left(p_4^2 - \sum_{i=1}^{3} p_i^2 - m_0^2 c^2 \right) \Psi_1 = 0, \qquad (22)$$

and one will similarly find second-order equations that relate to the other three Ψ. One has then verified a certain result in advance, since the matrices (17) satisfy the conditions (16).

The Dirac equations (18) present a very asymmetric aspect: The z-axis and the operator p_3 clearly play a special role in them.

The Ψ_i that are solutions to these equations are then initially coupled to the choice of axes as in Pauli's theory; they must be used in order to calculate the probabilities for which the z-axis plays a special role. If one takes other axes then one can write equations of propagation that still have the form (18), but they will have other functions Ψ_i for solutions that are coupled to the preceding ones by transformations that are analogous to the ones that we encountered in Pauli's theory; we shall examine that point in the following chapter.

There are some interesting remarks that can be made about the passage from the old equation (10) to equations (18). That passage is somewhat analogous to the passage from the equation for light waves to Maxwell's equations. Indeed, the equation of propagation of light waves (*in vacuo*) is:

$$\Delta u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0. \qquad (23)$$

It is a second-order equation in which u denotes a quantity that characterizes the light perturbation. In electromagnetic theory, u can be any of the six components of the two

fields of the light wave – viz., the electric field and the magnetic field. One will then have six equations of type (23) for the six quantities h_x, h_y, h_z, H_x, H_y, and H_z.

The passage from those equations of propagation to Maxwell's equations consists precisely of replacing the second-order equations with simultaneous first-order equations that couple the six quantities h_x, …, H_z in such a fashion that the six second-order equations of type (23) will be a consequence. That is indeed also the path that Dirac followed on order to pass from the second-order equations (15) to the first-order equations (15).

If the wave of wave mechanics has any physical reality in the classical sense then one would expect that the four Ψ_i would be the four components of a space-time vector. However, we know today that the wave of wave mechanics has no physical reality in the classical sense: It is a complex expression that serves only as an intermediary in the calculations and permits one to define certain real expressions that do have some physical meaning, such as the probability density $\Psi \Psi^*$. In Dirac's theory, the four Ψ_i do not by any means need to have the character of vector components, and we effectively see that they do not. However, there exist certain real combinations of these quantities that have a physical sense (as a probability) and possess a vectorial character. We expect to learn about them soon.

4. The Dirac equations in an electromagnetic field. – Equations (18) are suitable only in the case where a field is absent. What must one take to be the equations of propagation when the electron displaces in a field that is defined by a scalar potential V and a vector potential $\mathbf{A}$? Dirac answered that question by saying: It suffices to replace the operations p_i of the formulas (9) with the operators P_i of the formulas (6) in the symbolic equation (12).

If one assumes Dirac's postulate then the equations of propagation in an electromagnetic field can be written symbolically as:

$$(P_4 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 m_0 c)\, \Psi = 0, \tag{24}$$

or, explicitly:

$$(P_4 + m_0 c)\, \Psi_1 + (P_1 + i P_2)\, \Psi_4 + P_3 \Psi_3 = 0,$$

$$(P_4 + m_0 c)\, \Psi_2 + (P_1 - i P_2)\, \Psi_3 - P_3 \Psi_4 = 0,$$

$$(P_4 - m_0 c)\, \Psi_3 + (P_1 + i P_2)\, \Psi_2 + P_3 \Psi_1 = 0,$$

$$(P_4 - m_0 c)\, \Psi_4 + (P_1 - i P_2)\, \Psi_1 - P_3 \Psi_2 = 0.$$

$$\tag{25}$$

What is truly remarkable is that these equations, which are obtained by means of general considerations that are completely independent of the difficulties that were pointed out in the first part of this book contain the properties of the rotating, magnetic electron! We shall begin to understand this fact by seeking to define second-order equations that generalize equations (15) in the case where an electromagnetic field is present upon starting with equations (25).

In order to do that, we apply the operator $P_4 - \left(\sum_{i=1}^{3} \alpha_i P_i + \alpha_4 m_0 c \right)$ to the symbolic equation (24). We will then obtain:

$$\left(P_4 - \sum_{i=1}^{3} \alpha_i P_i - \alpha_4 m_0 c \right)\left(P_4 + \sum_{i=1}^{3} \alpha_i P_i + \alpha_4 m_0 c \right)\Psi = 0. \tag{26}$$

If we develop the indicated operators while taking the relations (16) into account then we will find that:

$$\left[P_4^{\ 2} + \sum_{i=1}^{3} \alpha_i (P_4 P_i - P_i P_4) - \sum_{i=1}^{3} P_i^{\ 2} - \sum_{i \neq j}(\alpha_i \alpha_j P_i P_j + \alpha_j \alpha_i P_j P_i) - m_0^{\ 2} c^2 \right]\Psi = 0. \tag{27}$$

We now recall that the electromagnetic fields are coupled to the potentials by the relations:

$$\mathbf{h} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{H} = \operatorname{rot} \mathbf{A}. \tag{28}$$

Upon taking into account the definitions (6) and the conditions (16), we will obtain, after a simple calculation:

$$\sum_{i=1}^{3} \alpha_i (P_4 P_i - P_i P_4) = -\frac{h}{2\pi i}\frac{e}{c}(\alpha_1 h_x + \alpha_2 h_y + \alpha_3 h_z),$$

$$\sum_{i \neq j}(\alpha_i \alpha_j P_i P_j + \alpha_j \alpha_i P_j P_i) = -\frac{h}{2\pi i}\frac{e}{c}(\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z). \tag{29}$$

Equation (27) will then become:

$$\left[P_4^{\ 2} - \sum_{i=1}^{3} P_i^{\ 2} - m_0^{\ 2} c^2 - \frac{e}{c}\cdot\frac{h}{2\pi i}(\alpha_1 h_x + \alpha_2 h_y + \alpha_3 h_z) \right.$$
$$\left. + \frac{e}{c}\cdot\frac{h}{2\pi i}(\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z) \right]\Psi = 0. \tag{30}$$

If just the first three terms in brackets exist then one will be back to equation (5), which will then be valid for each Ψ_i. The new element that is introduced by Dirac's theory is the intervention of the supplementary terms in (30). In order to discover the sense of those supplementary terms, look at the non-relativistic equation (2), which we can write in the form:

$$\left(2m_0 c \cdot p_4 - \sum_{i=1}^{3} p_i^2 - 2m_0 U \right) \Psi = 0. \tag{31}$$

Upon comparing (31) with (30), we perceive that we can consider the supplementary terms in question to be the potential energy terms, but on the condition that one must divide by $-2m_0$, except that the role that was played by m_0 in the previous theories will be played by $-\alpha_4 m_0$ in Dirac's mechanics, as we shall see later on ([23]). We then agree to define potential energy terms by *pre*-multiplying the two supplementary terms in (30) by $\dfrac{1}{2m_0 \alpha_4} = \dfrac{\alpha_4}{2m_0}$. Taking into account the commutation relations between the α_i, that will lead us to set:

$$U_e = - \frac{eh}{4\pi m_0 c} \, i \, (\alpha_1 \, \alpha_4 \, h_x + \alpha_2 \, \alpha_4 \, h_y + \alpha_3 \, \alpha_4 \, h_z),$$

$$U_m = - \frac{eh}{4\pi m_0 c} \, i \, (\alpha_2 \, \alpha_3 \, \alpha_4 \, H_x + \alpha_3 \, \alpha_1 \, \alpha_4 \, H_y + \alpha_1 \, \alpha_2 \, \alpha_4 \, H_z). \tag{32}$$

Now, one must remember that a body with an electric moment $\mathfrak{P}$ that is placed in an electric field **h** will possess a potential energy of $-(\mathfrak{P} \cdot \mathbf{h})$, while a body with a magnetic moment of $\mathfrak{M}$ that is placed in a magnetic field **H** will possess a potential energy of $-(\mathfrak{M} \cdot \mathbf{H})$. We are then led to attribute a magnetic moment to the electron whose components are:

$$\mathfrak{M}_x = \frac{eh}{4\pi m_0 c} i \, \alpha_2 \, \alpha_3 \, \alpha_4, \quad \mathfrak{M}_y = \frac{eh}{4\pi m_0 c} i \, \alpha_3 \, \alpha_1 \, \alpha_4, \quad \mathfrak{M}_z = \frac{eh}{4\pi m_0 c} i \, \alpha_1 \, \alpha_2 \, \alpha_4, \tag{33}$$

and an electric moment whose components are:

$$\mathfrak{P}_x = \frac{eh}{4\pi m_0 c} i \, \alpha_1 \, \alpha_4, \qquad \mathfrak{P}_y = \frac{eh}{4\pi m_0 c} i \, \alpha_2 \, \alpha_4, \qquad \mathfrak{P}_z = \frac{eh}{4\pi m_0 c} i \, \alpha_3 \, \alpha_4. \tag{34}$$

In order to speak more precisely, the quantities (33) and (34) are the *operators* that correspond to the components of the two moments. Since $\dfrac{eh}{4\pi m_0 c}$ is the Bohr magneton, one sees that the Dirac equations automatically attribute a proper magnetism to the electron that is coupled to the Bohr magneton. We then see the magnetism of the electron appear, and we will soon have occasion to make that first indication more precise. Moreover, formulas (34) show us that the Dirac electron also possesses an electric moment whose significance we shall see later on.

([23]) See Chapter XV, para. 4.

The components of the two moments that were defined by (33) and (34) are operators, since they are expressed with the aid of the α_i . That should not surprise us, since we are accustomed to seeing physical quantities give way to operators in the new mechanics. Moreover, it is easy to verify that the operators (33) and (34) are Hermitian, Indeed, since the α_i are Hermitian and anti-commute with each other, products such as $\alpha_1 \alpha_2$ or $\alpha_1 \alpha_2 \alpha_4$, for example, will be anti-Hermitian, and as a result, their product with i will be Hermitian. The operators (33) and (34) will then indeed have the desired character for something that would represent physical quantities.

We then find that we have eliminated a difficulty that existed in Dirac's original presentation, which wrote formulas (33) and (34) without the factor of α_4, and thus obtained non-Hermitian components for the electric moment.

_____

RELATIVISTIC INVARIANCE OF
THE DIRAC EQUATION

————

1. Invariance of the form of the Dirac equations under a Lorentz transformation.

The Dirac equations make the z-axis play a special role, and, as in Pauli's theory, the wave function serves to answer some questions of probability in which the z-axis is involved. If we would like to pose the same questions of probability for an axis oz' that is distinct from oz then we must write the Dirac equations in a system where oz' is the third axis. We must be able to write the Dirac equations *in the same form* for all systems of axes, while the four wave functions transform in a certain fashion when we pass from one system of axes to the other.

Not only do the Dirac equations, like the Pauli equations, possess just that invariance of form with respect to spatial coordinate changes, but they are also invariant in form under *all* Lorentz transformations, and thus satisfy the principle of relativity.

One knows that a Lorentz transformation of the most general type can always be decomposed into three successive transformations. Starting with the original system x, y, z, t of space-time coordinates, one performs:

1. A rotation around the oz axis, which is defined by formulas of the type:

$$x = x' \cos \alpha - y' \sin \alpha, \qquad y = y' \cos \alpha + x' \sin \alpha, \qquad z = z', \qquad t = t'. \quad (1)$$

2. A rotation around the oy axis, which is perpendicular to oz and is defined by:

$$z = z' \cos \theta - x' \sin \theta, \qquad x = x' \cos \alpha + z' \sin \alpha, \qquad y = y', \qquad t = t'. \quad (2)$$

3. A *simple* Lorentz transformation – i.e., the passage from one system $oxyzt$ to a system $o'x'y'z't'$ that is in uniform relative motion with respect to the first one, with the z and z' axes sliding over each other, while the other axes remain parallel, respectively. One will then get the well-known transformation formulas:

$$x = x', \qquad y = y', \qquad z = \frac{z' + \beta ct}{\sqrt{1 - \beta^2}}, \qquad t = \frac{t' + \dfrac{\beta}{c} x'}{\sqrt{1 - \beta^2}}, \qquad (3)$$

in which βc is the velocity of the second system with respect to the first one. If we set:

$$\frac{1}{\sqrt{1 - \beta^2}} = \cosh \gamma, \qquad \frac{\beta}{\sqrt{1 - \beta^2}} = \sqrt{\cosh^2 \gamma - 1} = \sinh \gamma \qquad (4)$$

then formulas (3) can be written:

$$x = x', \quad y = y', \quad z = z' \cosh \gamma + ct' \sinh \gamma, \quad ct = ct' \cosh \gamma + z' \sinh \gamma. \tag{5}$$

Upon combining three formulas of the types 1, 2, and 3, one can perform a Lorentz transformation of the most general type.

The invariance of form of the Dirac equations under a general Lorentz transformation will be proved if one can arrive at a solution of the following problem: Knowing that the Dirac equations:

$$(P_4 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 m_0 c) \, \Psi = 0 \tag{6}$$

admit the functions $\Psi_1 (x, y, z, t)$, $\Psi_2 (x, y, z, t)$, $\Psi_3 (x, y, z, t)$, $\Psi_4 (x, y, z, t)$ in a system of Galilean axes $xyzt$, show that in another system of Galilean axes $x'y'z't'$, the Dirac equations:

$$(P_4' + \alpha_1 P_1' + \alpha_2 P_2' + \alpha_3 P_3' + \alpha_1 m_0 c) \, \Psi' = 0, \tag{7}$$

in which, one has:

$$P_1' = -\frac{h}{2\pi i} \frac{\partial}{\partial x'} + \frac{e}{c} A_x', \quad \ldots, \quad P_4' = \frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t'} + \frac{e}{c} V', \tag{8}$$

will admit functions $\Psi_1'(x', y', z', t')$, $\Psi_2'(x', y', z', t')$, $\Psi_3'(x', y', z', t')$, and $\Psi_4'(x', y', z', t')$ for solutions that can be expressed linearly as functions of Ψ_1, ..., Ψ_4 by means of formulas that involve the parameters that define the passage from the system $oxyzt$ to the system $o'x'y'z't'$.

It suffices to solve the problem thus-posed for the three transformations 1, 2, and 3 that were pointed out above, since any Lorentz transformation can be decomposed into transformations of those three types.

1. A rotation around oz:

Start with the equation:

$$(P_4 + m_0 c) \, \Psi_1 + (P_1 + i P_2) \, \Psi_4 + P_3 \, \Psi_3 = 0,$$

$$(P_4 + m_0 c) \, \Psi_2 + (P_1 - i P_2) \, \Psi_3 - P_3 \, \Psi_4 = 0,$$

$$(P_4 - m_0 c) \, \Psi_3 + (P_1 + i P_2) \, \Psi_2 + P_3 \, \Psi_1 = 0, \tag{9}$$

$$(P_4 - m_0 c) \, \Psi_4 + (P_1 - i P_2) \, \Psi_1 - P_3 \, \Psi_2 = 0.$$

Since the change of variables is expressed by the formulas (1), one will have:

$$P_1 = P_1' \cos \alpha - P_2' \sin \alpha, \quad P_2 = P_2' \cos \alpha - P_1' \sin \alpha, \quad P_3 = P_3', P_4 = P_4', \tag{1'}$$

and upon substituting this in (9), one will easily find that:

$$(P'_4 + m_0 c)\Psi_1 + (P'_1 e^{i\alpha} + i P'_2 e^{i\alpha})\Psi_4 + P'_3 \Psi_3 = 0,$$

$$(P'_4 + m_0 c)\Psi_2 + (P'_1 e^{-i\alpha} - i P'_2 e^{-i\alpha})\Psi_3 - P'_3 \Psi_4 = 0,$$

$$(10)$$

$$(P'_4 - m_0 c)\Psi_3 + (P'_1 e^{i\alpha} + i P'_2 e^{i\alpha})\Psi_2 + P'_3 \Psi_1 = 0,$$

$$(P'_4 - m_0 c)\Psi_4 + (P'_1 e^{-i\alpha} - i P'_2 e^{-i\alpha})\Psi_1 - P'_3 \Psi_2 = 0.$$

In these equations, the Ψ_i are assumed to be expressed as functions of the primed variables with the aid of the transformation relations (1). Multiplying the first and third equations in (10) by $e^{-i\alpha/2}$, and second and fourth one by $e^{i\alpha/2}$, and adding them will give:

$$(P'_4 + m_0 c)\Psi_1 e^{-i\alpha/2} + (P'_1 + i P'_2)\Psi_4 e^{i\alpha/2} + P'_3 \Psi_3 e^{-i\alpha/2} \quad = 0,$$

$$(P'_4 + m_0 c)\Psi_1 e^{i\alpha/2} + (P'_1 - i P'_2)\Psi_3 e^{-i\alpha/2} - P'_3 \Psi_3 e^{i\alpha/2} \quad = 0,$$

$$(11)$$

$$(P'_4 - m_0 c)\Psi_1 e^{-i\alpha/2} + (P'_1 + i P'_2)\Psi_1 e^{-i\alpha/2} + P'_3 \Psi_3 e^{-i\alpha/2} = 0,$$

$$(P'_4 - m_0 c)\Psi_1 e^{i\alpha/2} + (P'_1 - i P'_2)\Psi_2 e^{i\alpha/2} - P'_3 \Psi_2 e^{i\alpha/2} \quad = 0.$$

The system (11) shows us that the functions Ψ'_i are coupled to the functions Ψ_i by the simple formulas:

$$\Psi'_1 (x', y', z', t') = \Psi_1 (x', y', z', t') \, e^{-i\alpha/2},$$

$$\Psi'_2 (x', y', z', t') = \Psi_2 (x', y', z', t') \, e^{i\alpha/2},$$

$$(12)$$

$$\Psi'_3 (x', y', z', t') = \Psi_3 (x', y', z', t') \, e^{-i\alpha/2},$$

$$\Psi'_4 (x', y', z', t') = \Psi_4 (x', y', z', t') \, e^{i\alpha/2},$$

and the required proof is found in this case.

2. Rotation around oy:

Since the transformation of coordinates is given by (2), the transformation of Dirac wave functions is given by the formulas:

$$\Psi'_1 (x', y', z', t') = \Psi_1 (x', y', z', t') \cos\frac{\theta}{2} + \Psi_2 (x', y', z', t') \sin\frac{\theta}{2},$$

$$\Psi'_2 (x', y', z', t') = \Psi_2 (x', y', z', t') \cos\frac{\theta}{2} - \Psi_1 (x', y', z', t') \sin\frac{\theta}{2},$$

$$\Psi'_3 (x', y', z', t') = \Psi_3 (x', y', z', t') \cos\frac{\theta}{2} + \Psi_4 (x', y', z', t') \sin\frac{\theta}{2}, \tag{13}$$

$$\Psi'_4 (x', y', z', t') = \Psi_4 (x', y', z', t') \cos\frac{\theta}{2} - \Psi_3 (x', y', z', t') \sin\frac{\theta}{2}.$$

One proves this by an argument that is analogous to the one that was made for case 1.

3. Simple Lorentz transformation:

Since the transformation of coordinates is expressed by the formulas (5), that of the Dirac function will be given by:

$$\Psi'_1 (x', y', z', t') = \Psi_1 (x', y', z', t') \cosh\frac{\gamma}{2} + \Psi_2 (x', y', z', t') \sinh\frac{\gamma}{2},$$

$$\Psi'_2 (x', y', z', t') = \Psi_2 (x', y', z', t') \cosh\frac{\gamma}{2} - \Psi_1 (x', y', z', t') \sinh\frac{\gamma}{2},$$

$$\tag{13}$$

$$\Psi'_3 (x', y', z', t') = \Psi_3 (x', y', z', t') \cosh\frac{\gamma}{2} + \Psi_4 (x', y', z', t') \sinh\frac{\gamma}{2},$$

$$\Psi'_4 (x', y', z', t') = \Psi_4 (x', y', z', t') \cosh\frac{\gamma}{2} - \Psi_3 (x', y', z', t') \sinh\frac{\gamma}{2}.$$

One again proves this in the same fashion, and the invariance in form of the Dirac equations for the most general Lorentz transformation is then proved.

One sees from formulas (12), (13), and (14) that the Ψ_i do not transform like coordinates. As we have said already, they do not have the character of components of a space-time vector. However, we hope to soon define certain expressions with the Ψ_i that do have a vectorial or tensorial character and are thus physically significant.

2. A more synthetic proof of relativistic invariance. – We shall indicate a more synthetic proof of the invariance of the Dirac equations. The proof that was developed by von Neumann is a generalization of the one that was employed by Pauli in order to prove the invariance of his equations with respect to the change of rectangular coordinates in space.

Take the Dirac equation in its symbolic form:

$$\left(P_4 + \sum_{i=1}^{3} \alpha_i P_i + \alpha_4 \, m_0 \, c \right) \Psi = 0, \tag{15}$$

and first apply the operator α_i . We will then get $(\alpha_4^2 = 1)$:

$$\left(\alpha_4 P_4 + \sum_{i=1}^{3} \alpha_4 \, \alpha_i P_i + m_0 \, c \right) \Psi = 0. \tag{16}$$

We shall now take the variables of the Minkowski universe:

$$x_1 = x, \qquad x_2 = y, \qquad x_3 = z, \qquad x_4 = ict, \tag{17}$$

in place of the variables x, y, z, t, and define the corresponding operators:

$$\pi_1 = -\frac{h}{2\pi i} \frac{\partial}{\partial x_i} + \frac{e}{c} A_1 = P_1, \qquad \pi_2 = P_2, \qquad \pi_3 = P_3,$$

$$\pi_4 = -\frac{h}{2\pi i} \frac{\partial}{\partial x_4} + \frac{e}{c} iV = -\frac{P_4}{i}. \tag{18}$$

Equation (16) is then written:

$$\left(\alpha_4 \, \pi_4 + \sum_{i=1}^{3} i\alpha_4 \, \alpha_i \, \pi_i + i \, m_0 \, c \right) \Psi = 0. \tag{19}$$

With von Neumann, we then set:

$$\gamma_1 = i \, \alpha_4 \, \alpha_1, \qquad \gamma_2 = i \, \alpha_4 \, \alpha_2, \qquad \gamma_3 = i \, \alpha_4 \, \alpha_3, \qquad \gamma_4 = \alpha_4. \tag{20}$$

It is easy to verify that one has:

$$\gamma_i^2 = 1, \qquad \gamma_i \, \gamma_j + \gamma_j \, \gamma_i = 0 \qquad (i \neq j), \tag{21}$$

which one can summarize by writing:

$$\gamma_i \, \gamma_j + \gamma_j \, \gamma_i = 2\delta_{ij} \cdot 1, \tag{22}$$

in which 1 is the identity matrix. One also verifies that the γ_i are Hermitian.

With those notations, equation (19) will take the condensed form that is used frequently:

$$\left(\sum_{i=1}^{4} \gamma_i \, \pi_i + i \, m_0 c \right) \Psi = 0. \tag{23}$$

Now suppose that we make a change of Galilean axes by way of a general Lorentz transformation. It is well-known in the theory of relativity that a general Lorentz transformation is equivalent to a rotation of the axes in the Minkowski universe. The new variables x_i' after the transformation will then be coupled to the old variables x_i by the formulas:

$$x_i = \sum_j o_{ij} x_j', \tag{24}$$

in which o is a matrix with four rows and four columns ([1]) that satisfy the orthogonality relation:

$$\sum_i o_{ik} o_{ij} = \delta_{kj}. \tag{25}$$

Now, it is obvious that the π_i transform like the x_i ; i.e., that:

$$\pi_i = \sum_i o_{ik} \pi_j'. \tag{26}$$

After the change of coordinates, equation (23) will then be written:

$$\left(\sum_{i=1}^4 \gamma_i \sum_{j=1}^4 o_{ij} \pi_j' + i m_0 c \right) \Psi = 0, \tag{27}$$

in which the functions Ψ must be expressed with the aid of new variables x_i'. If we set:

$$\gamma_j' = \sum_i o_{ij} \gamma_i \tag{28}$$

then we can replace (27) with:

$$\left(\sum_{i=1}^4 \gamma_i' \pi_j' + i m_0 c \right) \Psi = 0. \tag{29}$$

Equation (28) expresses the matrices γ_j' as functions of the matrices γ_j. If one would like to make its meaning more precise then one can write:

$$\gamma_{j,mn}' = \sum_{i=1}^4 o_{ij} \gamma_{i,mn}, \tag{30}$$

in which, for example, $\gamma_{i,mn}$ denotes the element of the matrix γ_i whose indices are m, n.
Since one has:

$$\gamma_{j,mn}'^* = \sum_{i=1}^4 o_{ij}^* \gamma_{i,mn}^* = \sum_{i=1}^4 o_{ij}^* \gamma_{i,mn} \neq \gamma_{j,mn}', \tag{31}$$

([1])   The matrix o is not real; those of its elements that contain the index 4 once will be pure imaginary.

by reason of the Hermiticity of the γ_i, one sees that, in general, the γ_i' are *not* Hermitian, because not all of the o_{ij} are real.

It is easy to verify that the γ_i' satisfy conditions of the form (22). In fact, one has:

$$\gamma_i' \gamma_j' + \gamma_j' \gamma_i' = \sum_k o_{ki} \gamma_k \cdot \sum_l o_{lj} \gamma_l + \sum_i o_{ij} \gamma_l \cdot \sum_k o_{ki} \gamma_k$$
$$= \sum_k \sum_l o_{ki} o_{lj} (\gamma_k \gamma_l + \gamma_l \gamma_k) \qquad (32)$$
$$= \sum_k \sum_l o_{ki} o_{lj} \cdot 2\delta_{kl} = 2\sum_k o_{ki} o_{kj} = 2\,\delta_{ij}.$$

The essential point of the argument is that there exists a matrix Λ with four rows and four columns such that one has:

$$\gamma_i' = \Lambda^{-1} \gamma_i \Lambda \qquad (i = 1, 2, 3, 4). \qquad (33)$$

If the relation (33) is true then the fact that the γ_i satisfy the relations (22) will imply that they satisfy the analogous relations:

$$\gamma_i' \gamma_j' + \gamma_j' \gamma_i' = 2\,\delta_{ij}. \qquad (34)$$

Equation (33) can be written:

$$\Lambda \gamma_i' = \gamma_i \Lambda \qquad (i = 1, 2, 3, 4), \qquad (35)$$

so in order to prove the *existence* of the matrix Λ, one must show that there exist 16 quantities Λ_{kl} that satisfy the 64 equations:

$$\sum_{l=1}^{4} \Lambda_{kl} \gamma_{i,lm}' = \sum_{l=1}^{4} \gamma_{i,kl} \Lambda_{lm}, \qquad (36)$$

in which the indices k, m, i can take on the values 1, 2, 3, 4. The existence of the matrix is then by no means obvious *a priori*. Meanwhile, we shall assume the existence of that matrix temporarily, and then justify that hypothesis *a posteriori*.

Before going any further, we remark that the matrix Λ cannot be unitary, in general. In fact, if it were unitary then we would have:

$$\Lambda^+ = \Lambda^{-1}, \qquad \Lambda = (\Lambda^{-1})^+, \qquad (37)$$

and as one will infer from (33) upon taking the adjoint equation:

$$\gamma_i'^+ = (\Lambda^{-1} \gamma_i \Lambda)^+ = \Lambda^+ \gamma_i^+ (\Lambda^{-1})^+, \qquad (38)$$

one will have:

$$\gamma_i'^+ = \Lambda^{-1} \gamma_i \Lambda = \gamma_i', \qquad (39)$$

by virtue of (37) and the Hermiticity of γ_i.

It will then result that γ_i' will be Hermitian, which is not generally true, as we have seen.

Now, recall equation (29) by introducing the relation (33). It will become:

$$\left(\sum_{j=1}^{4} \Lambda^{-1} \gamma_j \Lambda \, \pi_j' + i \, m_0 c \right) \Psi = 0. \tag{40}$$

Multiply this on the left by Λ, and upon remarking that the matrix Λ corresponds to an operation that is performed on the Dirac index, commute it with π_j'. One will get:

$$\left(\sum_{j=1}^{4} \gamma_j \, \pi_j' + i \, m_0 c \right) \Lambda \Psi = 0. \tag{41}$$

Formula (41) expresses the following theorem:

THEOREM: *When one makes a Lorentz transformation, one can keep the same form for the Dirac equations, but the new wave functions Ψ_i' will be related to the old ones Ψ_i by a linear transformation:*

$$\Psi_i'(x', y', z', t') = \Lambda \, \Psi_i(x', y', z', t') = \sum_{k=1}^{4} \Lambda_{ik} \Psi_k(x', y', z', t'). \tag{42}$$

That is precisely the result that was established already in paragraph 1.

We can now justify the hypothesis of the existence of the matrix Λ. Indeed, we learned in paragraph 1 how to calculate the linear transformation that each function Ψ_i is subjected to for each of the three types of coordinate change into which one can decompose any general Lorentz transformation. We then know (at least, in principle) how to calculate the linear transformation of the Ψ_i that corresponds to an arbitrary Lorentz transformation; i.e., to determine the elements of the matrix Λ such that:

$$\Psi_i'(x', y', z', t') = \Lambda \, \Psi_i(x', y', z', t'). \tag{43}$$

Now, from (42), that matrix Λ must be the same as the one that enters into (33), and since we know how to calculate it, we are sure that it exists.

Upon referring to formulas (12), (13), and (14), one can immediately write down the expressions for the matrices Λ that correspond to the simple cases 1, 2, 3, of the preceding paragraph. For example, consider the case 3: viz., a simple Lorentz transformation. Formulas (14) show us that the matrix Λ will then have the following form:

$$\Lambda = \begin{vmatrix} \cosh\dfrac{\gamma}{2} & 0 & \sinh\dfrac{\gamma}{2} & 0 \\ 0 & \cosh\dfrac{\gamma}{2} & 0 & -\sinh\dfrac{\gamma}{2} \\ \sinh\dfrac{\gamma}{2} & 0 & \cosh\dfrac{\gamma}{2} & 0 \\ 0 & -\sinh\dfrac{\gamma}{2} & 0 & \cosh\dfrac{\gamma}{2} \end{vmatrix}. \tag{44}$$

It is easy to verify that this matrix is not unitary ($\Lambda^+ \neq \Lambda^{-1}$). It will then result that $\sum\limits_{i=1}^{4} \Psi_i^* \Psi_i$ is not equal to $\sum\limits_{i=1}^{4} \Psi_i'^* \Psi_i'$. The probability density will then change in value for a simple Lorentz transformation. Indeed, we see that this density is not an invariant, but the temporal component of a space-time vector.

3. Electromagnetic invariance of the Dirac equations. – Aside from relativistic invariance, the Dirac equations present another type of invariance that I will call "electromagnetic invariance" (it is the "gauge" invariance, or *Eichinvarianz* of the German authors). Let me explain what that means.

Since the electric field **h** and the magnetic field **H** are defined by the formulas:

$$\mathbf{h} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{H} = \operatorname{rot} \mathbf{A}, \tag{45}$$

it is obvious that if one replaces *V* and **A** with:

$$V' = V - \frac{1}{c}\frac{\partial \Phi}{\partial t}, \qquad \mathbf{A}' = \mathbf{A} + \operatorname{grad} \Phi, \tag{46}$$

in which Φ is an arbitrary function of *x*, *y*, *z*, *t*, then one will not modify any of the fields, because one will have:

$$\operatorname{rot}\mathbf{A}' = \operatorname{rot}\mathbf{A}, \qquad -\operatorname{grad} V' - \frac{1}{c}\frac{\partial \mathbf{A}'}{\partial t} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}. \tag{47}$$

Since it is the fields that express the dynamical actions, and since they are insensitive to a transformation of the potentials of the form (46), we would have to expect from this that the Dirac equations would be invariant under those transformations (46). That is the electromagnetic invariance in question.

We write the symbolic Dirac equation:

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + \frac{e}{c}V \right) + \sum_{j=1}^{3} \alpha_j \left(-\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial x_j} + \frac{e}{c}A_j \right) + \alpha_4 m_0 c \right] \Psi = 0, \tag{48}$$

and then suppose that we have subjected the potentials to the transformation (46). Upon replacing V and A_j as functions of V' and A'_j, we will then get:

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{e}{c}V'+\frac{e}{c^2}\frac{\partial\Phi}{\partial t}\right)+\sum_{j=1}^{4}\alpha_j\left(-\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial x_j}+\frac{e}{c}A'_j-\frac{e}{c}\frac{\partial\Phi}{\partial x_j}\right)+\alpha_4\,m_0\,c\right]\Psi=0, \quad (49)$$

in place of (48).

It is easy to verify that if one sets $\Psi' = \Psi\,e^{-\frac{2\pi i\,e}{h\,c}\Phi}$ then one will recover an equation that has the same form as (48), namely:

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{e}{c}V'\right)+\sum_{j=1}^{4}\alpha_j\left(-\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial x_j}+\frac{e}{c}A'_j\right)+\alpha_4\,m_0\,c\right]\Psi'=0. \quad (49)$$

The passage from the potentials V and $\mathbf{A}$ to the potentials V' and $\mathbf{A}'$ does not modify the form of the Dirac equations then, but each of the Ψ_k will be multiplied by $e^{-\frac{2\pi i\,e}{h\,c}\Phi}$.

However, we know that it is not the Ψ_k that have any physical sense, but certain combinations of Ψ_k. Now, one can verify that any of the combinations of that type, which we hope to understand soon, will not be modified when one replaces Ψ_k with Ψ_k $e^{-\frac{2\pi i\,e}{h\,c}\Phi}$; that is immediately obvious for the probability density $\sum_{k=1}^{4}\Psi_k^*\,\Psi_k$, for example.

That insensitivity of quantities that are physically meaningful under the transformations (46) constitutes the electromagnetic invariance of the Dirac equations.

_____

DENSITY AND CURRENT IN DIRAC'S THEORY.
PLANE WAVES

———

1. Expressions for the probability density and current.

In Dirac's theory, we must seek to transpose the ideas of wave mechanics for a function Ψ. In particular, we must seek to define a probability density ρ of presence and a current $\rho \, \mathbf{u}$ for that probability. Here, the expressions $-e\rho$ and $-e\rho \, \mathbf{u}$ once more give the mean electric charge density and the mean electric current, with the aid of which, one calculates the mean radiation that is emitted by a set of electrons.

In order to find the form for ρ and $\rho \, \mathbf{u}$, we must always be guided by the idea that the total probability of presence must remain constant (viz., equal to 1), and that the equation of continuity $\partial \rho \, / \, \partial t + \operatorname{div} (\rho \, \mathbf{u}) = 0$ must be, in turn, a consequence of the equations of propagation.

We write the four Dirac equations and their four conjugates:

$$(P_4 + m_0 \, c) \, \Psi_1 + (P_1 + iP_2) \, \Psi_4 + P_3 \, \Psi_3 \ = 0,$$

$$(P_4 + m_0 \, c) \, \Psi_2 + (P_1 - iP_2) \, \Psi_3 - P_3 \, \Psi_4 \ = 0,$$

$$(P_4 - m_0 \, c) \, \Psi_3 + (P_1 + iP_2) \, \Psi_2 + P_3 \, \Psi_1 \ = 0,$$

$$(P_4 - m_0 \, c) \, \Psi_4 + (P_1 - iP_2) \, \Psi_1 - P_3 \, \Psi_2 \ = 0,$$

(1)

and

$$(P_4^* + m_0 c)\Psi_1^* + (P_1^* - i P_2^*)\Psi_4^* + P_3^* \Psi_3^* \ = 0,$$

$$(P_4^* + m_0 c)\Psi_2^* + (P_1^* + i P_2^*)\Psi_3^* - P_3^* \Psi_4^* \ = 0,$$

$$(P_4^* - m_0 c)\Psi_3^* + (P_1^* - i P_2^*)\Psi_2^* + P_3^* \Psi_1^* \ = 0,$$

$$(P_4^* - m_0 c)\Psi_4^* + (P_1^* + i P_2^*)\Psi_1^* - P_3^* \Psi_2^* \ = 0.$$

(2)

Multiply equations (1) by Ψ_1^*, Ψ_2^*, Ψ_3^*, Ψ_4^*, respectively, and equations (2) by Ψ_1, Ψ_2, Ψ_3, Ψ_4, respectively. Then take the sum of equations (1) and subtract the sum of equations (2); that will eliminate the terms in $m_0 \, c$. One will find terms of the form:

$$\Psi_i^* P_4 \Psi_i - \Psi_i P_4^* \, \Psi_i^* \ = \ \Psi_i^* \left(\frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} + \frac{e}{c} V \right) \Psi_i - \Psi_i \left(-\frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} + \frac{e}{c} V \right) \Psi_i^*$$

$$= \frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} (\Psi_i^* \Psi_i),$$

$$\Psi_i^* P_1 \Psi_i - \Psi_i P_1^* \Psi_i^* = \Psi_i^* \left(-\frac{h}{2\pi i} \frac{\partial}{\partial x} + \frac{e}{c} A_x \right) \Psi_i - \Psi_i \left(\frac{h}{2\pi i} \frac{\partial}{\partial x} + \frac{e}{c} A_x \right) \Psi_i^*$$

$$= -\frac{h}{2\pi i} \frac{\partial}{\partial x} (\Psi_i^* \Psi_i).$$

Finally, after multiplying by $\dfrac{2\pi i c}{h}$, one will find that:

$$\frac{\partial}{\partial t} \left(\Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4 \right)$$

$$+ \frac{\partial}{\partial x} \left[-c \left(\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1 \right) \right]$$

$$+ \frac{\partial}{\partial y} \left[-c \left(i\Psi_1^* \Psi_4 - i\Psi_2^* \Psi_3 + i\Psi_3^* \Psi_2 - i\Psi_4^* \Psi_1 \right) \right]$$

$$+ \frac{\partial}{\partial z} \left[-c \left(\Psi_1^* \Psi_3 - \Psi_2^* \Psi_4 + \Psi_3^* \Psi_1 - \Psi_4^* \Psi_2 \right) \right] = 0. \tag{3}$$

Equation (3) will be equivalent to the continuity equation if one sets:

$$\rho \quad = \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4,$$

$$\rho\, u_x = -c \left(\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1 \right),$$

$$\rho\, u_y = -c \left(i\Psi_1^* \Psi_4 - i\Psi_2^* \Psi_3 + i\Psi_3^* \Psi_2 - i\Psi_4^* \Psi_1 \right),$$

$$\rho\, u_y = -c \left(\Psi_1^* \Psi_3 - \Psi_2^* \Psi_4 + \Psi_3^* \Psi_1 - \Psi_4^* \Psi_2 \right). \tag{4}$$

One effortlessly verifies that the expressions (4) are real, because they are equal to their conjugates. One can then consider them to be the definition of the probability density of presence and the components of the corresponding current, respectively.

The expression for ρ indeed has the form that was postulated by Dirac. Moreover, it shows that one must normalize the Ψ_i by writing:

$$\int\limits_{-\infty}^{+\infty} \int \int (\Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4)\, d\tau = 1, \tag{5}$$

and that if that normalization is realized at an arbitrary epoch then it will always remain realized.

One can give a condensed and elegant form to the formulas (4) by appealing to the three matrices α_1, α_2, α_3, and the matrix 1 (with four rows and four columns). We write those matrices as:

$$\alpha_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 0 & 0 & 0 & +i \\ 0 & 0 & -i & 0 \\ 0 & +i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad 1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (6)$$

and compare them with formulas (4). We see forthwith that they can be written:

$$\rho = \sum_{i=1}^{4} \Psi_i^* \cdot 1 \Psi_i , \qquad\qquad \rho\, u_x = -c \sum_{i=1}^{4} \Psi_i^* \cdot \alpha_1 \Psi_i ,$$

$$\rho\, u_y = -c \sum_{i=1}^{4} \Psi_i^* \cdot \alpha_2 \Psi_i , \qquad \rho\, u_z = -c \sum_{i=1}^{4} \Psi_i^* \cdot \alpha_3 \Psi_i . \qquad (7)$$

Naturally, the mean electric charge density δ and the components of the mean current j_x, j_y, and j_z will be obtained upon multiplying ρ, $\rho\, u_x$, etc., by $-e$. One has:

$$\delta = -e\,\rho, \qquad j_x = -e\,\rho\, u_x , \quad j_y = -e\,\rho\, u_y , \quad j_z = -e\,\rho\, u_z . \qquad (8)$$

2. Vectorial character of the density and current. – The quantities (4), which have a physical sense, present the character of components of a space-time vector, with ρ as the temporal component. In order to prove that, it suffices, for example, to verify that the quantities (4) will transform like coordinates for each of the three transformations that were envisioned in the first paragraph of the preceding chapter, so it will then result that the same thing will be true for the most general Lorentz transformations.

We indicate the path to verification for a transformation of the type 1: viz., a rotation around oz. One will then have the following relations between the old and new variables:

$$x = x' \cos\alpha - y' \sin\alpha, \qquad y = y' \cos\alpha + x' \sin\alpha, \qquad z = z', \ \ t = t', \qquad (9)$$

so one will also have:

$$x' = x \cos\alpha + y \sin\alpha, \qquad y' = y \cos\alpha - x \sin\alpha, \qquad z' = z, \ \ t' = t. \qquad (9)$$

We have seen that the functions Ψ_i then transform in the following manner:

$$\Psi_1' = \Psi_1\, e^{-i\alpha/2}, \qquad \Psi_2' = \Psi_2\, e^{-i\alpha/2}, \qquad \Psi_3' = \Psi_3\, e^{-i\alpha/2}, \qquad \Psi_4' = \Psi_4\, e^{+i\alpha/2}. \quad (10)$$

It is then obvious that one will have:

$$\rho' = \sum_{i=1}^{4} \Psi_i'^* \, \Psi_i' = \sum_{i=1}^{4} \Psi_i^* \, \Psi_i = \rho. \qquad (12)$$

The density ρ remains invariant under the transformation, as the variable t does. One then sees just as easily that one will have $\rho' u'_z = \rho\, u_z$; i.e., the z component of the current will remain invariant like the variable z. For the x component, one will have:

$$
\begin{aligned}
\rho' u'_x &= - c\ (\Psi_1'^* \Psi_4' + \Psi_2'^* \Psi_3' + \Psi_3'^* \Psi_2' + \Psi_4'^* \Psi_1') \\
&= - c\ (\Psi_1^* \Psi_4\, e^{i\alpha} + \Psi_2^* \Psi_3\, e^{-i\alpha} + \Psi_3^* \Psi_2\, e^{i\alpha} + \Psi_4^* \Psi_1\, e^{-i\alpha}) \\
&= - c\ (\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1)\ \cos \alpha \\
&\quad - c\ (i\,\Psi_1^* \Psi_4 - i\,\Psi_2^* \Psi_3 + i\,\Psi_3^* \Psi_2 - i\,\Psi_4^* \Psi_1)\ \sin \alpha \\[2mm]
&= \rho\, u_x \cos \alpha + \rho\, u_y \sin \alpha.
\end{aligned}
\tag{13}
$$

The component $\rho\, u_x$ thus transforms like the variable x, and one can likewise verify that $\rho\, u_y$ transforms like y.

One proceeds in a similar way in order to study the transformations of the density and current in cases 2 and 3 (i. e., rotation around oy and simple Lorentz transformation, resp.), and one will arrive at the conclusion that the four quantities (4) are indeed the components of a space-time vector.

One knows that if a space-time vector **a** has spatial components a_1, a_2, a_3, and a temporal component a_4 then the quantity $c^2 a_4{}^2 - a_1{}^2 - a_2{}^2 - a_3{}^2$, which measures its length in space-time, will be an invariant that is independent of the chosen system of reference. If one calculates the length of the current density vector then one will find from a calculation that presents no difficulties (although it is a bit long) that:

$$
c^2 \rho^2 - (\rho\, u_x)^2 - (\rho\, u_y)^2 - (\rho\, u_z)^2 = c^2\, [\Omega_1{}^2 + \Omega_2{}^2],
\tag{14}
$$

with:

$$
\Omega_1 = \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 - \Psi_4^* \Psi_4 = \sum_{i=1}^{4} \Psi_i^* \cdot \alpha_4 \Psi_i\,,
$$

$$
\Omega_2 = + i\,\Psi_1^* \Psi_3 + i\,\Psi_2^* \Psi_4 - i\,\Psi_3^* \Psi_1 - i\,\Psi_4^* \Psi_3 = \sum_{i=1}^{4} \Psi_i^* \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_4 \Psi_i\,.
\tag{15}
$$

The quantities Ω_1 and Ω_2 are invariants, in such a way that the length (14) of current density vector in space-time is indeed invariant. The matrix $\alpha_1\, \alpha_2\, \alpha_3\, \alpha_4$ that is utilized in the condensed expression for Ω_2 is calculated easily by starting with the α_i ; it is Hermitian, and has the expression:

$$
\alpha_1\, \alpha_2\, \alpha_3\, \alpha_4 =
\begin{vmatrix}
0 & 0 & +i & 0 \\
0 & 0 & 0 & -i \\
+i & 0 & 0 & 0 \\
0 & +i & 0 & 0
\end{vmatrix}.
\tag{16}
$$

3. The plane wave in the absence of a field. – It is very instructive to treat the case in which a field is absent (viz., $V = \mathbf{A} = 0$). In that case, the Dirac equation will take the form:

$$\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+m_0 c\right)\Psi_1 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x}+i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_4 - \frac{h}{2\pi i}\frac{\partial\Psi_3}{\partial z} = 0,$$

$$\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+m_0 c\right)\Psi_2 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x}-i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_3 + \frac{h}{2\pi i}\frac{\partial\Psi_4}{\partial z} = 0,$$

$$\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}-m_0 c\right)\Psi_3 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x}+i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_2 - \frac{h}{2\pi i}\frac{\partial\Psi_1}{\partial z} = 0,$$

$$\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}-m_0 c\right)\Psi_4 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x}-i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_1 + \frac{h}{2\pi i}\frac{\partial\Psi_2}{\partial z} = 0.$$

(17)

Let us see if equations (17) admit the monochromatic plane wave solution that is defined by:

$$\Psi_i = a_i\, e^{\frac{2\pi i}{h}(Wt-p_x x-p_y y-p_z z)}.$$

(18)

Upon substituting this into (17), one will find that:

$$\left(\frac{W}{c}+m_0 c\right)a_1 + (p_x + i\,p_y)\,a_4 + p_z\,a_3 = 0,$$

$$\left(\frac{W}{c}+m_0 c\right)a_2 + (p_x - i\,p_y)\,a_3 - p_z\,a_4 = 0,$$

$$\left(\frac{W}{c}-m_0 c\right)a_3 + (p_x + i\,p_y)\,a_2 + p_z\,a_1 = 0,$$

$$\left(\frac{W}{c}-m_0 c\right)a_4 + (p_x - i\,p_y)\,a_1 - p_z\,a_2 = 0.$$

(19)

The homogeneous, linear equations can be satisfied simultaneously by non-zero a_i only if the determinant:

$$\begin{vmatrix} \dfrac{W}{c}+m_0c & 0 & p_z & p_x+ip_y \\[2.5em] 0 & \dfrac{W}{c}+m_0c & p_x-ip_y & -p_z \\[2.5em] p_z & p_x+ip_y & \dfrac{W}{c}-m_0c & 0 \\[2.5em] p_x-ip_y & -p_z & 0 & \dfrac{W}{c}-m_0c \end{vmatrix} \tag{20}$$

is zero. The calculation of this determinant, which is somewhat lengthy, shows that it is equal to:

$$\left(\frac{W^2}{c^2}-m_0^2c^2-p_x^2-p_y^2-p_z^2 \right)^2 .$$

It is zero if W, p_x, p_y, p_z are coupled by the well-known relation of relativistic mechanics:

$$\frac{W^2}{c^2}-p_x^2-p_y^2-p_z^2= m_0^2c^2 . \tag{21}$$

Suppose that relation is satisfied ([1]). Not only is the determinant (20) zero, but so are all of its first-order minors. We can then assign two of the four a_i arbitrarily. For example, we give two arbitrary values A and B to a_3 and a_4 , respectively. Equations (19) then determine a_1 and a_2 , and one will find that:

$$a_1 = - \frac{p_zA+(p_x+ip_y)B}{W/c+m_0c} , \qquad a_2 = - \frac{(p_x-ip_y)A-p_zB}{W/c+m_0c} . \tag{22}$$

One then sees that the plane wave (18) is defined entirely if one knows the amplitudes A and B.

One can make an interesting remark about formulas (22). In the new mechanics, as in the old relativistic mechanics, we can say that the Newtonian approximation is valid when the energy W is slightly greater than the rest energy $m_0 c^2$ (velocities that are small with respect to the velocity of light).

It results from (21) that if the Newtonian approximation is valid then each of the quantities p_x , p_y , p_z will be very small in comparison to $m_0 c$; we would then refer to formulas (22). When the Newtonian approximation is valid, their denominators will be reasonably equal to $2m_0 c$, and one will see that a_1 and a_2 are very small in comparison to A and B. It will then result that in the Newtonian approximation, the functions Ψ_1 and Ψ_2 will be almost negligible in comparison to Ψ_3 and Ψ_4 . In that case, one is then reduced

---

([1]) We assume, for the moment, that W is positive; i.e., that one has, from (21) that: $W / c = +\sqrt{m_0^2c^2+p_x^2+p_y^2+p_z^2}$. Later on (Chap. XX), we shall want to know what the negative solution $W / c = -\sqrt{m_0^2c^2+p_x^2+p_y^2+p_z^2}$ signifies.

to a problem with two Ψ_i, as in Pauli's non-relativistic theory. A particularly simple case in which the Newtonian approximation is found to be rigorously valid is that of one electron at rest in the system of reference that is employed. The functions Ψ_i are then:

$$\Psi_1 = \Psi_2 = 0, \qquad \Psi_3 = A \; e^{\frac{2\pi i}{h} m_0 c^2 t}, \qquad \Psi_4 = B \; e^{\frac{2\pi i}{h} m_0 c^2 t}. \tag{23}$$

Therefore, the waves Ψ_1 and Ψ_2 will be rigorously zero in a system of Galilean axes that are attached to the electron.

4. Density and current for a plane wave. – We just saw that in the absence of a field, the Dirac equations will admit the solution:

$$\Psi_1 = -\frac{p_z A + (p_x + i p_y) B}{W/c + m_0 c} \; e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)},$$

$$\Psi_2 = -\frac{(p_x - i p_y) A - p_z B}{W/c + m_0 c} \; e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)},$$

$$\Psi_3 = A \, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)}, \tag{24}$$

$$\Psi_4 = B \, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)},$$

in which the constants W, p_x, p_y, p_z are coupled by the relation (21). The constants A and B are arbitrary, except for the normalization condition.

We calculate the density ρ of the plane wave (24):

$$\rho = \sum_{i=1}^{4} \Psi_i^* \Psi_i = AA^* + BB^* + (AA^* + BB^*) \frac{p_x^2 + p_y^2 + p_z^2}{(W/c + m_0 c)^2}$$

$$= (AA^* + BB^*) \left(1 + \frac{W/c - m_0 c}{(W/c + m_0 c)^2}\right) = (AA^* + BB^*) \frac{2W/c}{W/c + m_0 c}.$$

We likewise calculate $\rho \, u_x$ from the definition in (4):

$$\rho \, u_x = - c \, (\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1) = (AA^* + BB^*) \frac{2 p_x c}{W/c + m_0 c}, \tag{26}$$

which one can write by taking (25) into account:

$$\rho\, u_x = \rho \cdot \frac{p_x c^2}{W}.$$ (27)

The component u_x of the velocity the probability of presence is then:

$$u_x = \frac{p_x c^2}{W}.$$ (28)

Now, the component v_x of the velocity of the electron in its classical conception is that:

$$p_x = \frac{m_0 c}{\sqrt{1-\beta^2}} = \frac{W}{c^2} v_x,$$ (29)

from which, it will result from a comparison with (28) that:

$$u_x = v_x.$$ (30)

Upon reasoning with the y and z components, one will obtain formulas that are analogous to (26), (27), and (28), and one can conclude that:

$$u_y = v_y, \qquad u_z = v_z.$$ (31)

Briefly: the velocity $\mathbf{u}$ of the probability in the plane wave is everywhere equal to the velocity $\mathbf{v}$ that old corpuscular concept would attribute to the electron that is associated with that plane wave.

Naturally, in Dirac's theory, as in wave mechanics with just one function Ψ, the plane wave will correspond to the case in which one knows the dynamical state of the corpuscle (i.e., p_x, p_y, p_z, and, in turn, W are known) exactly, but its position is totally unknown.

The expressions for the density of electricity and the density of electric current here are:

$$\delta = -e\rho, \qquad j_x = -e\,\rho\, v_x, \quad j_y = -e\,\rho\, v_y, \quad j_z = -e\,\rho\, v_z,$$ (32)

in which ρ has the value in (25).

_____

THE PROPER MAGNETISM OF THE ELECTRON

_____

1.   The "probability globule" in wave mechanics.

The original idea of Uhlenbeck and Goudsmit consisted of considering the electron as a small sphere of electricity in rotation around one of its diameters and, in turn, possessing (at least, in its proper system) a magnetic moment that is directed along that diameter. That concept cannot be preserved to the letter in the new mechanics, due to the impossibility of attributing a position and structure to the electron. Nevertheless, we shall see that with the aid of the fiction of a "probability fluid," it is possible to obtain a sort of mean image of the electron in Dirac's theory that approaches the image of Uhlenbeck and Goudsmit. In order to do that, we must first recall some points from the wave mechanics of one function Ψ.

When we imagine an electron motion that takes place on a grand scale – for example, the deviation of an electron by a magnetic field – it suffices for us to be able to describe that motion in the classical fashion, and thus, to attribute a localization to the electron that is compatible with the uncertainty relations. Now, it is easy to see in the formulas that under the usual experimental conditions, the wave length that is associated with the electron is much smaller than the smallest length that we can measure directly. It then results that it is possible to construct a small group of waves that is defined by the superposition of monochromatic plane waves of extremely close frequencies and whose dimensions are negligible at our scale. A precise observation that is made on an electron will thus permit us to attribute a state of motion and position to the electron that is well-defined in practice *at our scale*, without violating the uncertainty relations. The fictitious probability fluid – whose density is, by definition, equal to the intensity $\Psi\Psi^*$ – forms a sort of small globule in this case, in whose interior the applied force can be regarded as constant. Since the globule agrees reasonably well with its center of gravity, it will then result from Ehrenfest's theorem that the globule moves like a material point that obeys the laws of classical mechanics. Since the corpuscle can manifest its presence only in the interior of the globule, and the dimensions of the globule are negligible in practice, everything will take place as if the corpuscle itself obeyed the classical laws. That is how things happen in the macroscopic domain that is at the junction between the old and the new mechanics.

However, one must remark that the probability globule does not represent the internal structure of the electron, as one would be first led to think. The electric density – $e\rho$ in the interior of the globule does not represent a true electric density that exists in the interior of the electron, which one assumes to be extended. In the present theory, one assumes that the electron is point-like and the density – $e\rho$ is, as we have explained already, only a mean electric density. The probability globule is therefore only a sort of *mean* image of the possible localizations of the electron. It is that mean representation of

the classical conception of the electron that is approached most closely by the new mechanics. Moreover, it is upon studying the probability globule in the Dirac theory that we expect to see the proper magnetism of the electron appear in a form that is analogous to the original idea of Uhlenbeck and Goudsmit. We shall soon see that it is a good form.

We shall recall how one can obtain a simple model of the probability globule in the context of the wave mechanics of one function Ψ (Darwin). Suppose that the wave Ψ has the form:

$$\Psi(x, y, z, 0) = a \, e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \, e^{-\frac{2\pi i}{h} m_0 (v_x x + v_y y + v_z z)} \tag{1}$$

at the initial instant $t = 0$, where the amplitude, which has spherical symmetry around the origin of the coordinates, is a Gaussian function of the radius vector [1]. The amplitude becomes negligible once the distance from the origin becomes a small multiple of σ. One can say that the globule at the initial instant has dimensions of order of σ. The second exponential factor in (1) represents the phase factor of a plane wave at time zero. In order for the wave Ψ to be equivalent to a group of waves, the quantity σ that measures its dimensions must be small with respect to the wave length $h / m_0 v$. By reason of the extreme smallness of the latter quantity, σ can nevertheless be negligible at our scale.

Darwin studied the propagation of a globule of the form (1). He showed [2] that during a sufficiently short time interval the globule would be transported collectively with the velocity $\mathbf{v}$, in such a way that one would have:

$$\Psi(x, y, z, t) = a \, e^{-\frac{(x-v_x t)^2+(y-v_y t)^2+(z-v_z t)^2}{2\sigma^2}} \, e^{-\frac{2\pi i}{h}[Wt - m_0 (v_x x + v_y y + v_z z)]} \tag{2}$$

at time t.

This is in accord with Ehrenfest's theorem. However, the globule always has a tendency to spread while it evolves in time.

Without insisting upon thus latter point, we can say that Darwin's spherical globule provides us with a sort of image of the macroscopic motion of the electron. It is interesting to calculate the density and current of the probability that corresponds to it. In order to do that, we must appeal to some formulas from wave mechanics in one Ψ:

$$\rho = \Psi\Psi^*, \qquad \rho \, \mathbf{u} = \frac{h}{4\pi i \, m_0} [\Psi \, \mathrm{grad} \, \Psi^* - \Psi^* \, \mathrm{grad} \, \Psi]. \tag{3}$$

One easily finds that:

$$\rho = a^2 \, e^{-\frac{x^2+y^2+z^2}{\sigma^2}}, \qquad\qquad \rho \, \mathbf{u} = \Psi\Psi^* \, \mathbf{v} = \rho \, \mathbf{v}, \tag{4}$$

at the initial instant, so:

$$\mathbf{u} = \mathbf{v}.$$

---

[1] In order for Ψ to be normalized, one must have $[a] = \pi^{-3/4} \, \sigma^{-3/2}$.

[2] One will find the details of the calculations in the author's book: *Introduction à l'étude de la Mécanique ondulatoire.* Paris, Hermann, 1930, chapter XIII.

The motion of the probability (or of the mean electric distribution) is therefore a convection with velocity **v**, and the Darwin globule indeed gives a mean macroscopic image of the classical electron from that viewpoint. Upon transposing the Darwin globule to Dirac's theory, we shall see Uhlenbeck and Goudsmit's spinning, magnetic electron appear.

2. The spherical probability globule in Dirac's theory. – Now, let us pass on to Dirac's wave mechanics with 4 functions Ψ. We have seen that one can give the initial amplitudes of Ψ_3 and Ψ_4 arbitrarily, and that Ψ_1 and Ψ_2 will then be deduced from the equations of propagation. We then give Ψ_1 and Ψ_2 a form here that was inspired by (1):

$$\Psi_3(x, y, z, 0) = A\ e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\ e^{-\frac{2\pi i}{h}m_0(v_x x+v_y y+v_z z)},$$

$$\Psi_4(x, y, z, 0) = B\,e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\ e^{-\frac{2\pi i}{h}m_0(v_x x+v_y y+v_z z)}.$$

(6)

Suppose that the Newtonian approximation is valid: We saw what that signifies in the last chapter. We can then replace $\dfrac{h}{2\pi i}\dfrac{\partial \Psi_3}{\partial t}$ and $\dfrac{h}{2\pi i}\dfrac{\partial \Psi_4}{\partial t}$ with $m_0 c^2$, and first two Dirac equations will give:

$$\left.\begin{aligned}
\Psi_1 &= \frac{1}{2m_0 c}\frac{h}{2\pi i}\left[\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial x}\right)\Psi_4+\frac{\partial \Psi_3}{\partial z}\right], \\
\Psi_2 &= \frac{1}{2m_0 c}\frac{h}{2\pi i}\left[\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial x}\right)\Psi_3-\frac{\partial \Psi_4}{\partial z}\right].
\end{aligned}\right\}$$

(7)

Set:

$$P = e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\ e^{-\frac{2\pi i}{h}m_0(v_x x+v_y y+v_z z)}.$$

(8)

We find the values of the derivatives of Ψ_3 and Ψ_4 upon substituting in (7):

$$\Psi_1(x, y, z, 0) = -\frac{1}{2m_0 c}\left[A\left(m_0 v_z+\frac{h}{2\pi i}\frac{z}{\sigma^2}\right)+B\left(m_0(v_x+iv_y)+\frac{h}{2\pi i}\frac{x+iy}{\sigma^2}\right)\right]P,$$

$$\Psi_2(x, y, z, 0) = -\frac{1}{2m_0 c}\left[A\left(m_0(v_x+iv_y)+\frac{h}{2\pi i}\frac{x+iy}{\sigma^2}\right)-B\left(m_0 v_z+\frac{h}{2\pi i}\frac{z}{\sigma^2}\right)\right]P.$$

(9)

These formulas give Ψ_1 and Ψ_2 at the instant 0 for a small value of v/c.

The functions Ψ_1 and Ψ_2 are almost negligible in comparison to Ψ_3 and Ψ_4, as we predicted.

If we form the expression for the density $\rho = \sum\limits_{i=1}^{4} \Psi_i^* \, \Psi_i$ then the first two terms in the sum Σ can be neglected in comparison to the last two, and it will suffice to write:

$$\rho = (AA^* + BB^*) \, e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} . \tag{10}$$

Naturally, one must have (normalization):

$$(AA^* + BB^*) \iint_{-\infty}^{+\infty}\!\!\int e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \, dx \, dy \, dz = 0. \tag{11}$$

Upon multiplying (10) by $-e$, one will obtain the mean electric density δ.

Naturally, one must calculate the components $\rho \, \mathbf{u}$ now. Here, none of the terms will be negligible, since the four terms in the expressions for these quantities will be products of wave functions of small value with wave functions of large value. One will have, for example:

$$\rho \, u_x = -c \, (\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1), \tag{12}$$

which will give, from (7) and (9):

$$\rho \, u_x = \left[AA^*\left(v_x - \frac{h}{2\pi m_0} \frac{y}{\sigma^2} \right) + BB^*\left(v_x + \frac{h}{2\pi m_0} \frac{y}{\sigma^2} \right) + \frac{h}{2\pi i \, m_0} \frac{z}{\sigma^2}(AB^* - BA^*) \right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

$$= \rho v_x + \frac{h}{2\pi m_0}\left[(AA^* - BB^*)\frac{y}{\sigma^2} - \frac{z}{\sigma^2}(iAB^* - iBA^*) \right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}},$$

so:

$$\rho \, u_x = \rho v_x + \frac{h}{4\pi m_0}\left[(AA^* - BB^*)\frac{\partial}{\partial y} - (iA^*B - iAB^*)\frac{\partial}{\partial z} \right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}. \tag{13}$$

One likewise finds:

$$\rho \, u_y = -c \, (-i\Psi_1^* \Psi_4 + i\Psi_2^* \Psi_3 - i\Psi_3^* \Psi_2 + i\Psi_4^* \Psi_1)$$

$$= \rho v_y + \frac{h}{4\pi m_0}\left[(AB^* + BA^*)\frac{\partial}{\partial z} - (AA^* - BB^*)\frac{\partial}{\partial x} \right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}},$$

$$\tag{13^{bis}}$$

$$\rho \, u_z = -c \, (\Psi_1^* \Psi_3 - \Psi_2^* \Psi_4 + \Psi_3^* \Psi_1 - \Psi_4^* \Psi_1)$$

$$= \rho v_z + \frac{h}{4\pi m_0}\left[(iA^*B - iAB^*)\frac{\partial}{\partial x} - (AB^* + A^*B)\frac{\partial}{\partial y} \right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}.$$

In order to obtain the components of the density of mean electric current, it will suffice to multiply ρu_x, ρu_y, ρu_z by $- e$, or more precisely, in order to obtain the expression in e.s.u., multiply them by $- e / c$. One will then find that:

$$j_x = -\frac{e}{c}\rho v_x + \frac{eh}{4\pi m_0 c}\left[(BB^* - AA^*)\frac{\partial}{\partial y} - (iAB^* - iA^*B)\frac{\partial}{\partial z}\right]e^{-\frac{x^2+y^2+z^2}{\sigma^2}},$$

$$j_y = -\frac{e}{c}\rho v_y + \frac{eh}{4\pi m_0 c}\left[-(A^*B + AB^*)\frac{\partial}{\partial z} - (BB^* - A^*A)\frac{\partial}{\partial x}\right]e^{-\frac{x^2+y^2+z^2}{\sigma^2}}, \tag{14}$$

$$j_z = -\frac{e}{c}\rho v_z + \frac{eh}{4\pi m_0 c}\left[(iAB^* - iA^*B)\frac{\partial}{\partial x} - (A^*B - AB^*)\frac{\partial}{\partial y}\right]e^{-\frac{x^2+y^2+z^2}{\sigma^2}}.$$

These formulas must now be interpreted.

3. Proof of a formula from electromagnetism. – In order to arrive at an interpretation of formulas (14), we shall have need for a formula from electromagnetism that we shall prove in this paragraph.

One knows that the magnetic action of a permanent current of density **j** is defined by the vector potential:

$$\mathbf{A} = \iiint \frac{\mathbf{j}\,d\tau}{r} \qquad (d\tau = \text{volume element}), \tag{15}$$

and the magnetic field is deduced from **A** by the formula:

$$\mathbf{H} = \text{rot } \mathbf{A} . \tag{16}$$

If one considers the electric current to be composed of a sheaf of electrons in motion, and if δ denotes the mean density of the charge in that sheaf of electrons then one must set $\mathbf{j} = \delta \mathbf{v}$, where **v** is the velocity of the electrons, which is assumed to be uniform. However, if the electrons are endowed with a magnetic moment then the expression for **j** must be completed by taking into account the intensity of the magnetization **I** that then exists in the sheaf. We shall examine how **j** depends upon **I**.

Let a small magnet that is composed of two magnetic masses $- \mu$ and $- \mu$ be situated at a distance l from each other. If α, β, γ are the direction cosines of the axis of the same magnet then the magnetic moment **m** of the magnet will have the components:

$$m_x = \alpha\mu l = \alpha m, \qquad m_y = \beta m, \qquad m_z = \gamma m. \tag{17}$$

The magnetic potential χ that is created by the magnet at a point M that is situated at a distance r from it center is:

$$\chi = \frac{\mu}{r_1} - \frac{\mu}{r_2} = \mu\left(\frac{\partial \frac{1}{r}}{\partial x}\alpha l + \frac{\partial \frac{1}{r}}{\partial y}\beta l + + \frac{\partial \frac{1}{r}}{\partial z}\gamma l\right) = \left(\mathbf{m}\cdot\mathrm{grad}\,\frac{1}{r}\right), \tag{18}$$
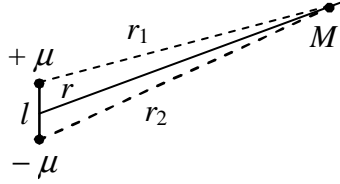
Figure 6.

and the field that is created at M will have the components:

$$H_x = -\frac{\partial \chi}{\partial x} = -\left(m_x\frac{\partial^2 \frac{1}{r}}{\partial x^2} + m_y\frac{\partial^2 \frac{1}{r}}{\partial x\,\partial y} + m_z\frac{\partial^2 \frac{1}{r}}{\partial x\,\partial z}\right),$$

$$H_y = -\frac{\partial \chi}{\partial y} = -\left(m_x\frac{\partial^2 \frac{1}{r}}{\partial y\,\partial x} + m_y\frac{\partial^2 \frac{1}{r}}{\partial y^2} + m_z\frac{\partial^2 \frac{1}{r}}{\partial y\,\partial z}\right), \tag{19}$$

$$H_z = -\frac{\partial \chi}{\partial z} = -\left(m_x\frac{\partial^2 \frac{1}{r}}{\partial z\,\partial x} + m_y\frac{\partial^2 \frac{1}{r}}{\partial z\,\partial y} + m_z\frac{\partial^2 \frac{1}{r}}{\partial z^2}\right).$$

We shall show that the vector potential $\mathbf{A}$ − whose $\mathbf{H}$ field, as defined by (19), must be its rotation − is given by:

$$\mathbf{A} = \left[\mathbf{m}\cdot\mathrm{grad}\,\frac{1}{r}\right], \tag{20}$$

in which the brackets indicate an exterior product. Indeed, one has for H_x :

$$H_x = (\mathrm{rot}\,\mathbf{A})_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = \frac{\partial}{\partial y}\left(m_x\frac{\partial \frac{1}{r}}{\partial y} - m_y\frac{\partial \frac{1}{r}}{\partial x}\right) - \frac{\partial}{\partial z}\left(m_z\frac{\partial \frac{1}{r}}{\partial x} - m_x\frac{\partial \frac{1}{r}}{\partial z}\right)$$

$$= -m_y\frac{\partial^2 \frac{1}{r}}{\partial y\,\partial x} - m_z\frac{\partial^2 \frac{1}{r}}{\partial z\,\partial x} + m_x\left[\frac{\partial^2 \frac{1}{r}}{\partial y^2} + \frac{\partial^2 \frac{1}{r}}{\partial z^2}\right], \tag{21}$$

so, by virtue of the well-known relation $\Delta\frac{1}{r} = 0$, one will infer the first expression (19)

for H_x . One will likewise find the expressions (19) for H_y and H_z . This is, indeed, the vector potential (20) that corresponds to the magnetic field (19).

Now, suppose that we are concerned, not with a small magnet of moment **m**, but with an extended magnetic body whose magnetization **I** we know at each point. We must replace **m** with $\mathbf{I}\,d\tau$ in the latter formulas and integrate, which will give:

$$\mathbf{A} = \iiint \left[\mathbf{I} \cdot \operatorname{grad} \frac{1}{r} \right] d\tau. \tag{22}$$

Suppose that the vector **I** is zero on the boundary of the magnetic body. An integration by parts will then permit us to write:

$$A_x = \iiint \left[I_y \frac{\partial \frac{1}{r}}{\partial z} - I_z \frac{\partial \frac{1}{r}}{\partial y} \right] d\tau = \iiint_V \frac{1}{r} \left(\frac{\partial I_z}{\partial y} - \frac{\partial I_y}{\partial z} \right) d\tau = \iiint_V \frac{(\operatorname{rot} \mathbf{I})_x}{r} \, d\tau, \tag{23}$$

(where V is the volume of the magnetic body) and we will have analogous formulas for A_y and A_z. We will thus have the vectorial relation:

$$\mathbf{A} = \iiint_V \frac{\operatorname{rot} \mathbf{I}}{r} \, d\tau. \tag{24}$$

Finally, if we are dealing with a body that is both electrified and magnetized and in uniform motion with velocity **v** then the vector potential that is created by that body will be, by virtue of (15) and (24):

$$\mathbf{A} = \iiint_V \frac{\delta \mathbf{v} + \operatorname{rot} \mathbf{I}}{r} \, d\tau. \tag{25}$$

Similarly, it will then be created by a current of density:

$$\mathbf{j} = \delta \mathbf{v} + \operatorname{rot} \mathbf{I}. \tag{26}$$

It is formula (26) that must serve for the interpretation of formulas (14).

4. Interpretation of formulas (14). – Return to formulas (14), and define a vector **I** by giving the following values to its components:

$$I_x = \frac{eh}{4\pi m_0 c} (-A^* B - AB^*) e^{-\frac{x^2 + y^2 + z^2}{\sigma^2}},$$

$$I_y = \frac{eh}{4\pi m_0 c} (iAB^* - iA^* B) e^{-\frac{x^2 + y^2 + z^2}{\sigma^2}}, \tag{27}$$

$$I_z = \frac{eh}{4\pi m_0 c}(BB^* - AA^*)e^{-\frac{x^2+y^2+z^2}{\sigma^2}} .$$

The vector **I** is real (because $I_x = I_x^*$, etc.) and it is zero on the boundary of the probability globule; i.e., at infinity, here. With that vector **I**, formulas (14) can be written vectorially:

$$\mathbf{j} = -\frac{e}{c}\rho\,\mathbf{v} + \text{rot}\,\mathbf{I} . \tag{28}$$

Formula (28) has great interest to us because upon comparing it with formula (26), we see that the probability globule – which is a mean macroscopic representation of the electron – cannot be assimilated into a simple sphere of charge – e that moves with a velocity of **v**, but into a sphere that is both electrified and magnetized whose intensity of magnetization is equal to **I** at each of its points.

The total magnetic moment of the globule is the vector $\mathfrak{M}$ that one obtains by integrating the vector **I**. Taking (27) into account and the normalization condition (11), one finds that the components of $\mathfrak{M}$ are ([1]):

$$\mathfrak{M}_x = \iiint I_x d\tau = \frac{eh}{4\pi m_0 c}(-A^*B - AB^*)\iiint e^{-\frac{x^2+y^2+z^2}{\sigma^2}} = -\frac{eh}{4\pi m_0 c}\frac{A^*B + AB^*}{AA^* + BB^*},$$

$$\mathfrak{M}_y = \iiint I_y d\tau = \frac{eh}{4\pi m_0 c}\frac{i(AB^* - A^*B)}{AA^* + BB^*}, \tag{29}$$

$$\mathfrak{M}_z = \iiint I_z d\tau = \frac{eh}{4\pi m_0 c}\frac{BB^* - AA^*}{AA^* + BB^*}.$$

The length of the vector $\mathfrak{M}$ is then:

$$|\mathfrak{M}| = \sqrt{\mathfrak{M}_x^2 + \mathfrak{M}_y^2 + \mathfrak{M}_z^2}$$

$$= \frac{eh}{4\pi m_0 c}\sqrt{\frac{(A^*B + AB^*)^2 - (A^*B - AB^*)^2 + (BB^* - AA^*)^2}{(AA^* + BB^*)^2}}$$

$$= \frac{eh}{4\pi m_0 c} . \tag{30}$$

([1]) In reality, $\mathfrak{M}_x$, $\mathfrak{M}_y$, $\mathfrak{M}_z$ are only the mean values, as we shall indicate in a precise manner in the following chapter. One must then write $\bar{\mathfrak{M}}_x$, instead of $\mathfrak{M}_x$, etc.

The globule thus has a magnetic moment that is equal to a Bohr magneton.

Now, recall formula (13), and seek to write it in the form: $\rho\, u_x = \rho\, v_x + \rho\, v'_x$. We are thus led to set:

$$
\begin{aligned}
v'_x &= \frac{h}{2\pi m_0 \rho}\left[(BB^* - AA^*)\frac{y}{\sigma^2} - (iAB^* - iA^*B)\frac{z}{\sigma^2}\right]e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \\
&= \frac{h}{2\pi m_0 \sigma^2}\left[\frac{BB^* - AA^*}{AA^* + BB^*}\,y - \frac{iAB^* - iA^*B}{AA^* + BB^*}\,z\right],
\end{aligned}
\tag{31}
$$

upon taking (10) into account. We likewise set:

$$
v'_y = \frac{h}{2\pi m_0 \sigma^2}\left[\frac{-(AB^* + A^*B)}{AA^* + BB^*}\,z - \frac{(BB^* - AA^*)}{AA^* + BB^*}\,x\right],
$$

$$
\tag{31, cont.}
$$

$$
v'_z = \frac{h}{2\pi m_0 \sigma^2}\left[\frac{iAB^* - iA^*B}{AA^* + BB^*}\,x - \frac{-(AB^* + A^*B)}{AA^* + BB^*}\,y\right],
$$

and then set:

$$
\omega_x = \frac{h}{2\pi m_0 \sigma^2}\frac{(AB^* + A^*B)}{AA^* + BB^*}, \qquad \omega_y = \frac{-h}{2\pi m_0 \sigma^2}\frac{iAB^* - iA^*B}{AA^* + BB^*},
$$

$$
\omega_z = \frac{-h}{2\pi m_0 \sigma^2}\frac{BB^* - AA^*}{AA^* + BB^*}.
\tag{32}
$$

Formulas (31) and (31, cont.) then become:

$$
v'_x = \omega_y z - y\,\omega_z, \qquad v'_y = \omega_z x - z\,\omega_x, \qquad v'_z = \omega_x y - x\,\omega_y.
\tag{33}
$$

Since we have $\mathbf{u} = \mathbf{v} + \mathbf{v}'$, from the manner itself by which we have introduced $\mathbf{v}$, we will see that the velocity of the probability is the sum of the velocity $\mathbf{v}$ of translation of an ensemble of the globule and an *internal* velocity $\mathbf{v}'$ that is due to a rotation of the ensemble that is defined by the velocity of rotation $\boldsymbol{\omega}$ whose components are ω_x, ω_y, ω_z. That velocity of rotation will be represented by the vector $\boldsymbol{\omega}$, which passes through the center of the globule and is parallel to $\mathfrak{M}$ (since one will have $\omega_x : \omega_y : \omega_z = \mathfrak{M}_x : \mathfrak{M}_y : \mathfrak{M}_z$) and has a length:

$$
|\boldsymbol{\omega}| = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2} = \frac{h}{2\pi m_0 \sigma^2}.
\tag{34}
$$

This internal rotation of the ensemble of the mean electric fluid explains the origin of the magnetic moment $\mathfrak{M}$. It becomes faster as the globule becomes smaller, which is easily explained, since the magnetic moment $\mathfrak{M}$ must always be equal to one Bohr magneton. The probability globule in Dirac's theory thus gives a sort of mean macroscopic image of the spinning, magnetic electron.

_____

THE "MAGNETIC AND ELECTRIC MOMENT" TENSOR DENSITY

———

1. The magnetic moment of the Dirac electron in the Newtonian approximation.

In the last chapter, while studying the spherical probability globule that was defined by THE wave functions (6) and (9), we were led to exhibit a vector **I** that was given by formulas (27), which was a vector that represented the magnetization intensity − i.e., the magnetic moment density – in the globule. Now, if one takes formulas (6) into account then one can write formulas (27), which give the components of **I**, in the form:

$$I_x = \frac{eh}{4\pi m_0 c}(-\Psi_3^* \Psi_4 - \Psi_3 \Psi_4^*),$$

$$I_y = \frac{eh}{4\pi m_0 c}(i\Psi_3 \Psi_4^* - i\Psi_4 \Psi_3^*), \tag{1}$$

$$I_z = \frac{eh}{4\pi m_0 c}(\Psi_4^* \Psi_4 - \Psi_3^* \Psi_3).$$

In this new form, the expressions for the components of the vector **I** are valid for any Dirac wave in the Newtonian approximation (i.e., when Ψ_1 and Ψ_2 are negligible in comparison to Ψ_3 and Ψ_4), and not just for the spherical probability globule that was envisioned in the last chapter.

The components of the mean magnetic moment of the Dirac electron are then obtained by integrating the expressions (1) over all of space. One will then have:

$$\bar{\mathfrak{M}}_x = \iiint I_x \, d\tau = \frac{eh}{4\pi m_0 c}\iiint (-\Psi_3^* \Psi_4 - \Psi_3 \Psi_4^*)\, d\tau,$$

$$\bar{\mathfrak{M}}_y = \iiint I_y \, d\tau = \frac{eh}{4\pi m_0 c}\iiint (i\Psi_3 \Psi_4^* - i\Psi_4 \Psi_3^*)\, d\tau, \tag{2}$$

$$\bar{\mathfrak{M}}_z = \iiint I_z \, d\tau = \frac{eh}{4\pi m_0 c}\iiint (\Psi_4^* \Psi_4 - \Psi_3^* \Psi_3)\, d\tau.$$

One must not forget that these formulas all have only a statistical significance. If one considers a very large number of electrons that are all found in the same state and are defined by the same functions Ψ_3 and Ψ_4 , and if one measures, for example, the

component of the proper magnetic moment along the *x*-axis for each of these electrons then one will find different results according to the special cases, but the mean value of the result that is obtained for the ensemble will be the $\overline{\mathfrak{M}}_x$ of formulas (2). That is only an application of the general ideas of new mechanics. In particular, one recalls that by virtue of a remark that was made in Chapter VI, paragraph 3, the expressions (1) are not true physical densities in the old sense, but are the quantities that one must integrate in order to obtain the mean values (2).

The third formula in (2) must merit some of our attention. We know that if one measures the proper magnetic moment of the electron parallel to the *z*-axis then one must necessarily find ± a Bohr magneton. The Dirac equations must play a role that is specific to the *z*-axis precisely, because the probabilities of these two hypotheses must be expressed simply with the aid of the Ψ_i. If one then casts one's eyes upon formula (2) then one will see that the probability that one must find a value of $+\dfrac{eh}{4\pi m_0 c}$ for $\mathfrak{M}_z$ is

$\iiint \Psi_4^* \Psi_4 \, d\tau$, whereas that of finding $-\dfrac{eh}{4\pi m_0 c}$ is $\iiint \Psi_3^* \Psi_3 \, d\tau$. This is in perfect accord with Pauli's ideas, which is what Dirac's theory reduces to when one assumes the Newtonian approximation here. The mean value (2) of $\mathfrak{M}_z$ will indeed represent the mean result of the measurement of $\mathfrak{M}_z$ for a large number of electrons in the state that is defined by Ψ_3 and Ψ_4.

The expressions for the components of **I** – i.e., the expressions (1) – are invariant in form for every change of coordinates in space. That must say that if one passes from one rectangular coordinate system *x*, *y*, *z* to another rectangular system *x′*, *y′*, *z′* then the components of **I** in the new system will be expressed with the aid of the new wave functions Ψ_i' in the same way that the old components of **I** were expressed by (1) with the aid of the Ψ_i. We shall not verify this here, since it is easy.

However, that invariance in form is no longer true for a Lorentz transformation. One thus notes that the expressions (1) have the character of the Newtonian approximation. In order to obtain relativistic invariance, one must take Ψ_1 and Ψ_2 into account and insert the three components of the magnetization intensity **I** into the six components of an anti-symmetric tensor of order two, as we shall see.

2. Mean magnetic moment of a plane wave in the Newtonian approximation. – Consider a plane wave that is defined in the Newtonian approximation by the two wave functions:

$$\Psi_3 = A \ e^{\frac{2\pi i}{h}[Wt - m_0 v_x x - m_0 v_y y - m_0 v_z z]} \ ,$$

$$\Psi_4 = B \ e^{\frac{2\pi i}{h}[Wt - m_0 v_x x - m_0 v_y y - m_0 v_z z]} \ .$$

$$(3)$$

What are the components of the mean magnetic moment?

With the aid of formulas (2), and upon taking the normalization into account ([1]), one will easily find that:

$$\overline{\mathfrak{M}}_x = \frac{eh}{4\pi m_0 c} \cdot \frac{-A^*B - AB^*}{AA^* + BB^*},$$

$$\overline{\mathfrak{M}}_y = \frac{eh}{4\pi m_0 c} \cdot \frac{iAB^* + iA^*B}{AA^* + BB^*}, \tag{4}$$

$$\overline{\mathfrak{M}}_z = \frac{eh}{4\pi m_0 c} \cdot \frac{BB^* - AA^*}{AA^* + BB^*}.$$
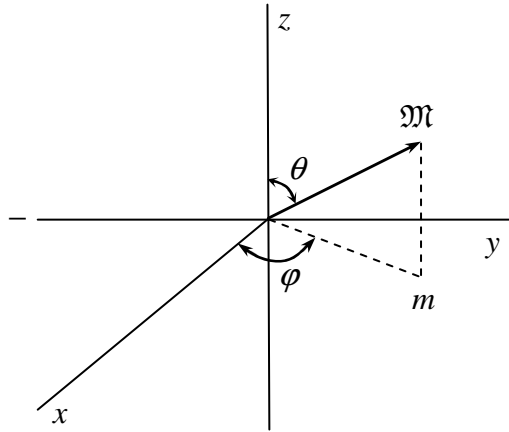
Figure 7.

Formulas (4) coincide, moreover, with formulas (29) of the last chapter. We could expect that to be true since the aforementioned formulas (29) are valid for any value of σ in formulas (6), and if we make σ tend to infinity then the spherical globule in the last chapter will tend to the plane wave (3).

Refer the direction of the vector $\mathfrak{M}$ to a spherical coordinate system θ and φ.

We will then have:

$$\overline{\mathfrak{M}}_x : \overline{\mathfrak{M}}_y : \overline{\mathfrak{M}}_z = (-A^*B - AB^*) : (iAB^* - iA^*B) : (BB^* - AA^*)$$

$$= \sin\theta\cos\varphi : \sin\theta\sin\varphi : \cos\theta, \tag{5}$$

so one easily infers that:

([1]) One might hesitate to write the normalization condition since the integration domain is infinite, and rigorously speaking one must introduce proper differentials. However, one can get around the difficulty in practice by first supposing that the domain is finite and has volume V. The normalization condition here can then be written:

$$\iiint (|\Psi_3|^2 + |\Psi_4|^2)\, d\tau = (AA^* + BB^*)\, V = 1,$$

and one will infer from (2), for example:

$$\overline{\mathfrak{M}}_x = \frac{eh}{4\pi m_0 c}(-A^*B - B^*A)\, V = \frac{eh}{4\pi m_0 c}\frac{-A^*B - B^*A}{AA^* + BB^*}.$$

Since this result is exact, no matter how large V is, formulas (4) will be found to be justified.

$$\left(-\frac{B}{A} - \frac{B^*}{A^*} \right) : i \left(-\frac{B}{A} + \frac{B^*}{A^*} \right) : \left(\frac{B}{A} \cdot \frac{B^*}{A^*} - 1 \right)$$

$$= 2 \cot \frac{\theta}{2} \cos \varphi : 2 \cot \frac{\theta}{2} \sin \varphi : \left(\cot^2 \frac{\theta}{2} - 1 \right). \tag{6}$$

Now, one satisfies equations (6) by setting:

$$-\frac{B}{A} = \cot \frac{\theta}{2} \, e^{i\varphi}, \tag{7}$$

and the complex relation (7), which is equivalent to two real relations, will show us how the orientation of $\mathfrak{M}$ is coupled to the value of the ratio B / A.

One can express this coupling in the following fashion (Darwin, Jordan): Consider a sphere of unit radius. The orientation of the vector $\mathfrak{M}$ is defined by the coordinates θ and φ of the point M at which that vector pierces the sphere. One projects that point M stereographically onto the plane of the equator while the center of the projection is the North pole.

The projection point m has the coordinates:

$$x = \cot \frac{\theta}{2} \cos \varphi, \qquad\qquad y = \cot \frac{\theta}{2} \sin \varphi. \tag{8}$$
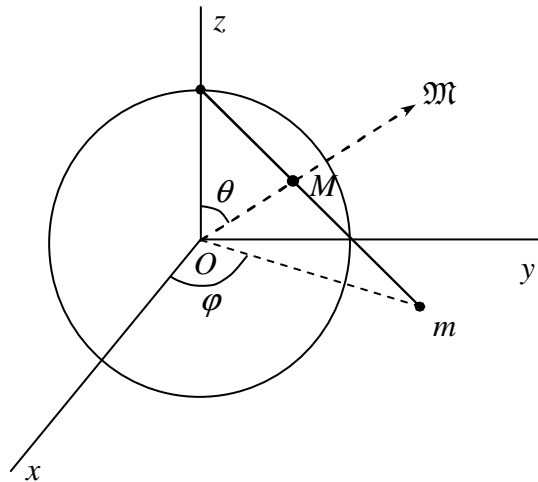

Figure 8.

If one considers the xy-plane to the plane of a complex variable then the number $\cot g \frac{\theta}{2} e^{i\varphi}$ will be affixed to the point m and, from (7), it will be equal to $- B / A$. One will then have the following result: The ratio $- B / A$ is coupled to the direction of the magnetic moment by the same relationship that couples the number that is affixed to a

point in the plane of a complex variable with the direction that it corresponds to on the sphere by stereographic projection.

3. The "magnetic and electric moment" tensor density. – We have seen that the expressions (1) possess invariance in form for coordinate changes in space, but not necessarily for Lorentz transformations. It is nevertheless possible to find six quadratic combinations of the four functions Ψ_i that transform under Galilean coordinate changes like the components of an anti-symmetric tensor of order two. Here are their expressions, which are all real:

$$\mu_{yz} = I_x = \frac{eh}{4\pi m_0 c} (\Psi_1^* \Psi_2 + \Psi_2^* \Psi_1 - \Psi_3^* \Psi_4 - \Psi_4^* \Psi_3) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_2 \alpha_3 \alpha_4 \Psi_k,$$

$$\mu_{zx} = I_y = \frac{eh}{4\pi m_0 c} (i\Psi_1^* \Psi_2 - i\Psi_2^* \Psi_1 - i\Psi_3^* \Psi_4 + i\Psi_4^* \Psi_3) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_3 \alpha_1 \alpha_4 \Psi_k,$$

$$\mu_{xy} = I_z = \frac{eh}{4\pi m_0 c} (\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_1 \alpha_2 \alpha_4 \Psi_k,$$

$$\mu_{xt} = J_x = \frac{eh}{4\pi m_0 c} (i\Psi_1^* \Psi_4 + i\Psi_2^* \Psi_3 - i\Psi_3^* \Psi_2 - i\Psi_4^* \Psi_1) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_1 \alpha_4 \Psi_k,$$ (9)

$$\mu_{yt} = J_y = \frac{eh}{4\pi m_0 c} (-\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 - \Psi_4^* \Psi_1) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_2 \alpha_4 \Psi_k,$$

$$\mu_{zt} = J_z = \frac{eh}{4\pi m_0 c} (i\Psi_1^* \Psi_3 - i\Psi_2^* \Psi_4 - i\Psi_3^* \Psi_1 + i\Psi_4^* \Psi_2) = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_3 \alpha_4 \Psi_k.$$

With the aid of the six quantities (9), and upon posing the antisymmetry relations $\mu_{zy} = -\mu_{yz}$, etc., one can define an antisymmetric table of four rows and four columns:

$$\begin{vmatrix} 0 & \mu_{xy} & \mu_{xz} & \mu_{xt} \\ \mu_{yx} & 0 & \mu_{yz} & \mu_{yt} \\ \mu_{zx} & \mu_{zy} & 0 & \mu_{zt} \\ \mu_{tx} & \mu_{ty} & \mu_{tz} & 0 \end{vmatrix}.$$ (10)

It is easy to verify that the table (10) defines an antisymmetric tensor of order two. In order to do this, it will suffice to verify that the μ_{ij} will transform according to the following schema:

$$\mu'_{ij} = \sum_{k,l} \mu_{kl} \frac{dx'_i}{dx_k} \frac{dx'_j}{dx_l} \qquad (11)$$

for each of the three simple transformations into which one can decompose a general Lorentz transformation.

If one can neglect the wave functions Ψ_1 and Ψ_2, which will be true when the Newtonian approximation is valid, then the quantities μ_{yz}, μ_{zx}, and μ_{xy} will reduce to I_x, I_y, and I_z, respectively, by formula (1). It is therefore natural to think that these three components of μ will be the three components of the mean magnetization $\mathbf{I}$ in their exact relativistic expressions.

What do the three components μ_{xt}, μ_{yt}, μ_{zt} then represent? In the theory of relativity, one must combine the magnetic field and the electric field in order to define an anti-symmetric tensor of order two. In a memoir on the magnetic electron before Dirac's theory, Frenkel showed the necessity of completing the magnetic moment of the electron with an electric moment ([1]). Here is how Frenkel reasoned: In the old theory of the magnetic electron, the electron was a corpuscle that possessed a proper magnetic moment. In a reference system in which the electron is in motion, it must create an electric field around it that is due to its magnetic moment because just as an electric charge in motion is equivalent to a current, and will thus creates a magnetic field around itself, a magnetic pole in motion will create an electric field around itself. The magnetic electron in motion must therefore possess an electric moment, and in order to satisfy the demands of the principle of relativity, the three components of the electric moment must be combined with the three components of the magnetic moment in order to permit us to define an antisymmetric tensor of order two. We will then be led to consider μ_{xt}, μ_{yt}, μ_{zt} to be the three components of an electric moment density.

The six quantities (9) permit us to form two invariants; i.e., there exist two combinations of the μ_{ij} whose value is the same in all reference systems. These two combinations are the following ones:

$$I^2 - J^2 = I_x^2 + I_{xy}^2 + I_z^2 - J_x^2 - J_y^2 - J_z^2 , \qquad (\mathbf{I} \cdot \mathbf{J}) = I_x J_x + I_y J_y + I_z J_z . \qquad (12)$$

Indeed, if we introduce the two invariants that were encountered already in Chapter XII, paragraph 3 then:

$$\Omega_1 = \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 - \Psi_4^* \Psi_4 = \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_4 \Psi_k ,$$

$$(13)$$

$$\Omega_2 = i\Psi_1^* \Psi_3 + i\Psi_2^* \Psi_4 - i\Psi_3^* \Psi_1 - i\Psi_4^* \Psi_2 = \sum_{k=1}^{4} \Psi_k^* \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_4 \Psi_k .$$

One easily verifies the relations:

([1]) Zeitschrift für Physik, 37, 4-5, pp. 243.

$$I^2 - J^2 = \left(\frac{eh}{4\pi m_0 c}\right)^2 (\Omega_1^2 - \Omega_2^2), \qquad (\mathbf{I} \cdot \mathbf{J}) = \left(\frac{eh}{4\pi m_0 c}\right)^2 \Omega_1 \Omega_2, \qquad (14)$$

which indeed shows the invariance of the two quantities. That invariance is, moreover, obvious from the tensorial viewpoint, since on the one hand, $I^2 - J^2$ is the "length" of the tensor in space-time, and on the other hand, the scalar product:

$$(\mathbf{I} \cdot \mathbf{J}) = \mu_{yz}\, \mu_{xt} + \mu_{zx}\, \mu_{yt} + \mu_{xy}\, \mu_{zt},$$

is obviously an invariant under coordinate changes in space, but no longer for a simple Lorentz transformation, as is easy to verify. Its invariance under a general Lorentz transformation will result from that.

As before, we can remark that the expressions (9) are not physical densities in the old sense of the word: They are only quantities that one must integrate in space in order to obtain the mean values of the components of the magnetic and electric moments of the electron. These two mean moments, which we denote by $\mathfrak{M}$ and $\mathfrak{P}$, will be given by the formulas:

$$\mathfrak{M} = \iiint \mathbf{I}\, d\tau, \qquad \mathfrak{P} = \iiint \mathbf{J}\, d\tau, \qquad (15)$$

in which the vectors $\mathbf{I}$ and $\mathbf{J}$ are defined by their component in (9). Naturally, the values of the components $\overline{\mathfrak{M}}_z$, etc., must be interpreted statistically. The formula:

$$\overline{\mathfrak{M}}_z = \iiint I_z\, d\tau = \frac{eh}{4\pi m_0 c} \iiint (\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4)\, d\tau \qquad (16)$$

signifies that a measurement of $\mathfrak{M}_z$ can provide the value $+\dfrac{eh}{4\pi m_0 c}$ with a probability of $\iiint (\Psi_1^* \Psi_1 + \Psi_4^* \Psi_4)\, d\tau$ and the value $-\dfrac{eh}{4\pi m_0 c}$, with a probability of:

$$\iiint (\Psi_2^* \Psi_2 + \Psi_3^* \Psi_3)\, d\tau.$$

The sum of the two probabilities is equal to unity by virtue of the normalization condition on Ψ_k.

Frenkel ([1]), by arguing with the old, purely corpuscular image of the electron, showed that the magnetic moment $\mathfrak{M}$ of the electron in a system where its velocity is $\mathbf{v}$ must be coupled to its electric moment $\mathfrak{P}$ in the same system by the relation:

([1]) *Loc. cit.*

$$\mathfrak{P} = \left[\mathfrak{M} \cdot \frac{\mathbf{v}}{c} \right]. \tag{17}$$

We shall see, in an example, that the Frenkel condition remains exact in Dirac's theory.

4. Simple example: the monochromatic plane wave. – We appeal to the formulas of Chapter XII, paragraph 3, and simply take the direction of propagation of the plane wave to be the z-axis. We will then have $p_x = p_x = 0$, and the four Ψ_i will have the form:

$$\Psi_3 = A\ e^{\frac{2\pi i}{h}(Wt - p_z z)}, \qquad\qquad \Psi_4 = B\ e^{\frac{2\pi i}{h}(Wt - p_z z)},$$

$$\Psi_1 = -\ \frac{p_z A}{\dfrac{W}{c} + m_0 c}\ e^{\frac{2\pi i}{h}(Wt - p_z z)}, \qquad\qquad \Psi_2 = \frac{p_z A}{\dfrac{W}{c} + m_0 c}\ e^{\frac{2\pi i}{h}(Wt - p_z z)}. \tag{18}$$

With that form for Ψ_i, formulas (9) are given easily:

$$I_x = -\ \frac{eh}{4\pi m_0 c}(A^* B + AB^*)\left(1 + \frac{p_z^2}{\left(\dfrac{W}{c} + m_0 c \right)} \right),$$

$$J_x = \frac{eh}{4\pi m_0 c}\ \frac{2ip_z}{\dfrac{W}{c} + m_0 c}(AB^* - A^* B),$$

$$I_y = i\frac{eh}{4\pi m_0 c}(AB^* - BA^*)\left(1 + \frac{p_z^2}{\left(\dfrac{W}{c} + m_0 c \right)} \right),$$

$$J_y = \frac{eh}{4\pi m_0 c}\ \frac{2ip_z}{\dfrac{W}{c} + m_0 c}(A^* B + AB^*), \tag{19}$$

$$I_z = -\ \frac{eh}{4\pi m_0 c}(BB^* - AA^*)\left(1 + \frac{p_z^2}{\left(\dfrac{W}{c} + m_0 c \right)} \right), \qquad J_x = 0.$$

One verifies painlessly that one has:

$$(\mathbf{I} \cdot \mathbf{J}) = 0. \tag{20}$$

This is, moreover, a consequence of the second formula in (14), since the invariant Ω_2 will be zero in the present case ([1]).

The examination of formulas (19) provides the relations:

$$J_x = I_y \frac{2 p_z \left(\dfrac{W}{c} + m_0 c \right)}{\left(\dfrac{W}{c} + m_0 c \right)^2 + p_z^2} = I_y \frac{v_z}{c} ,$$

$$\tag{21}$$

$$J_y = - I_x \frac{2 p_z \left(\dfrac{W}{c} + m_0 c \right)}{\left(\dfrac{W}{c} + m_0 c \right)^2 + p_z^2} = - I_x \frac{v_z}{c} , \qquad J_z = 0.$$

Upon integrating the relations (21) over all of space, in order to make the components of $\mathfrak{M}$ and $\mathfrak{P}$ appear, one sees that the Frenkel relation (17) is verified (since $v_x = v_y = 0$).

When the Newtonian approximation is valid, the ratio $\dfrac{p_z}{W/c + m_0 c}$ will be reasonably equal to $\dfrac{m_0 v_z}{2 m_0 c} = \dfrac{1}{2} \dfrac{v_z}{c}$, and its square will be negligible compared to unity. One will then find the values for I_x, I_y, I_z that one obtained in the preceding chapter upon starting with formulas (27) by letting σ tend to infinity, which is a result that one could have expected.

One can make the following remarks about the general formulas (19) and (20): The vector $\mathbf{J}$ is normal to the plane that is determined by the vector $\mathbf{I}$ and the vector $\mathbf{p}$. If the speed tends towards the speed of light – i.e., if p_z tends towards W / c (which is then much greater than $m_0 c$) – then the trihedron that is composed of the three vectors $\mathbf{p}$, $\mathbf{I}$, and $\mathbf{J}$ will tend to become a tri-rectangle, while the lengths of $\mathbf{I}$ and $\mathbf{J}$ will tend to become equal.

––––––––––––

([1]) Indeed, one can obviously calculate that invariant in no particular Galilean system – for example, in a system where the velocity of the electron is zero. Now, in such a system, one will have $\Psi_1 = \Psi_2 = 0$, and (13) will give $\Omega_2 = 0$.

CHAPTER XV

MATRICES AND FIRST INTEGRALS
IN DIRAC'S THEORY.
PROPER ANGULAR MOMENTUM OF THE ELECTRON.

———

1. The proper values and proper functions of the Dirac equations.

The proper values and proper functions of the Dirac equations are defined easily by analogy with the situation in the theory of a single function Ψ.

When the external field is independent of time, there will exist monochromatic solutions of the Dirac equations; i.e., solutions for which the four Ψ_k depend upon time only by the same exponential factor $e^{\frac{2\pi i}{h}Wt}$. The four Ψ_k will then satisfy the equations:

$$\left[\left(\frac{W}{c}+\frac{e}{c}V\right)+\alpha_1 P_1 +\alpha_2 P_2 +\alpha_3 P_3 +\alpha_4 m_0 c\right]\Psi_k = 0 \qquad (k = 1, 2, 3, 4). \qquad (1)$$

The values of the constant W for which there exists at least one set of Ψ_k that are finite, uniform, continuous, and zero at infinity are the "proper values" of equations (1). Any proper value W_n will then correspond to at least one set ([1]) of proper functions $\Psi_{1,n}$, $\Psi_{2,n}$, $\Psi_{3,n}$, and $\Psi_{4,n}$. We give two indices to each Ψ_k, such that the first one is the index that is introduced by Dirac's theory, and the second one characterizes the corresponding proper value.

By definition, the four functions $\Psi_{k,n}$ that correspond to the proper values W_n will then satisfy the equations:

$$\left[\left(\frac{W_n}{c}+\frac{e}{c}V\right)+\sum_{i=1}^{3}\alpha_i P_i +\alpha_4 m_0 c\right]\Psi_{k,n} = 0 \qquad (k = 1, 2, 3, 4). \qquad (2)$$

Since one can multiply the four functions $\Psi_{k,n}$ by the same arbitrary constant without them ceasing to satisfy equations (2), one will see that each proper solution is defined only up to a complex multiplicative constant. One determines the modulus of that constant with the aid of the normalization condition, which, as we have seen, must be written:

$$\int \sum_{k=1}^{4}\Psi_{k,n}^{*}\Psi_{k,n}\, d\tau = 1. \qquad (3)$$

We shall now show that:

([1]) The set of four proper functions $\Psi_{k,n}$ can be called a "proper solution" of equation (1).

1. The proper values W_n are all real.

2. If the proper functions $\Psi_{k,n}$ and $\Psi_{k,m}$ correspond to distinct proper values W_n and W_m then one will have the orthogonality condition:

$$\int \sum_{k=1}^{4} \Psi^*_{k,m} \Psi_{k,n} \, d\tau = 0. \tag{4}$$

Indeed, since the functions $\Psi_{k,n}$ obey equations (2), the functions $\Psi^*_{k,m}$ will obey the equations:

$$\left[\left(\frac{W_n^*}{c} + \frac{e}{c}V\right) + \sum_{i=1}^{3} \alpha_i^* P_i^* + \alpha_4^* m_0 c\right]\Psi^*_{k,m} = 0 \qquad (k = 1, 2, 3, 4). \tag{5}$$

Multiply (2) by $\Psi^*_{k,m}$ and (5) by $\Psi_{k,n}$, and then sum each equation over the index k and subtract them; one will get:

$$\sum_{k=1}^{4}\left[\frac{W_n - W_m^*}{c}\Psi^*_{k,m}\Psi_{k,n} + \Psi^*_{k,m}\sum_{i=1}^{3}\alpha_i P_i \cdot \Psi_{k,n} - \Psi_{k,n}\sum_{i=1}^{3}\alpha_i^* P_i^* \cdot \Psi^*_{k,m}\right.$$
$$\left. + m_0\, c\,(\Psi^*_{k,m}\alpha_4\Psi_{k,n} - \Psi_{k,n}\alpha_4^*\Psi^*_{k,m})\right] = 0. \tag{6}$$

We shall now show that if F is a linear operator that operates on the coordinates (and not on the Dirac indices k) then one will have:

$$\sum_{k=1}^{4}\Psi^*_{k,n} \cdot F\alpha_i\Psi_{k,n} = \sum_{k=1}^{4}F\Psi_{k,n} \cdot \alpha_i^* \Psi^*_{k,n} \qquad (i = 1, 2, 3). \tag{7}$$

Indeed, upon taking into account the Hermiticity of the α_i, one will find that:

$$\sum_{k=1}^{4}\Psi^*_{k,n} \cdot F\alpha_i\Psi_{k,n}$$
$$= \sum_{k=1}^{4}\Psi^*_{k,n} \cdot F\sum_{j=1}^{4}(\alpha_i)_{k,j}\Psi_{j,n} = \sum_{j,k=1}^{4}\Psi^*_{k,n}(\alpha_i)_{k,j}F\Psi_{j,n}$$
$$= \sum_{j,k=1}^{4}\Psi^*_{k,n}(\alpha_i^*)_{k,j}F\Psi_{j,n} = \sum_{j=1}^{4}F\Psi_{j,n}\cdot\sum_{k=1}^{4}(\alpha_i^*)_{j,k}\Psi^*_{k,n}$$
$$= \sum_{j=1}^{4}F\Psi^*_{j,n}\cdot\alpha_i^*\Psi_{j,n} \qquad \text{Q. E. D.} \tag{8}$$

We first infer a special case of (7) by setting $F = 1$:

$$\sum_{k=1}^{4} \Psi_{k,n}^{*} \cdot \alpha_i \Psi_{k,n} = \sum_{k=1}^{4} \Psi_{k,n} \cdot \alpha_i^{*} \Psi_{k,n}^{*} . \tag{9}$$

It will then result that the terms in $e A_x$, $e A_y$, $a A_z$, and $m_0 c$ disappear from (6). Equation (6) will then reduce to:

$$\sum_{k=1}^{4} \left[\frac{W_n - W_m^{*}}{c} \Psi_{k,m}^{*} \Psi_{k,n} - \frac{h}{2\pi i} \Psi_{k,m}^{*} \left(\frac{\partial}{\partial x} \alpha_1 + \frac{\partial}{\partial y} \alpha_2 + \frac{\partial}{\partial z} \alpha_3 \right) \Psi_{k,n} \right.$$
$$\left. - \frac{h}{2\pi i} \Psi_{k,n} \left(\frac{\partial}{\partial x} \alpha_1^{*} + \frac{\partial}{\partial y} \alpha_2^{*} + \frac{\partial}{\partial z} \alpha_3^{*} \right) \Psi_{k,m}^{*} \right] = 0. \tag{10}$$

Now, from (7), one will have:

$$\sum_{k=1}^{4} \Psi_{k,n}^{*} \frac{\partial}{\partial x} \alpha_i \Psi_{k,n} = \sum_{k=1}^{4} \frac{\partial \Psi_{k,n}}{\partial x} \cdot \alpha_i^{*} \Psi_{k,n}^{*} , \qquad \text{etc.,} \tag{11}$$

and formula (10) will become:

$$\sum_{k=1}^{4} \left[\frac{W_n - W_m^{*}}{c} \Psi_{k,m}^{*} \cdot \Psi_{k,n} \right.$$
$$\left. - \frac{h}{2\pi i} \left\{ \frac{\partial}{\partial x} (\Psi_{k,n} \alpha_1^{*} \Psi_{k,m}^{*}) + \frac{\partial}{\partial y} (\Psi_{k,n} \alpha_2^{*} \Psi_{k,m}^{*}) + \frac{\partial}{\partial z} (\Psi_{k,n} \alpha_3^{*} \Psi_{k,m}^{*}) \right\} \right] = 0. \tag{12}$$

Upon integrating over all of space, the term in curly brackets will give zero, because the Ψ_k are zero at infinity, and what will remain is:

$$\frac{W_n - W_m^{*}}{c} \int \sum_{k=1}^{4} \Psi_{k,m}^{*} \Psi_{k,n} \, d\tau = 0. \tag{13}$$

If one next sets $n = m$ then one will deduce from (13) that $W_n^{*} = W_n$ for any value of n, and that will signify that all of the W_n are real. If one next takes $n \neq m$, while remembering that $W_n \neq W_m$, by hypothesis, then one will deduce the orthogonality formula (4) from (13). The theorem that was just stated will then be proved.

The orthogonality relation generally breaks down for two proper solutions that correspond to the same proper values (viz., the degenerate case). However, the proper solutions that correspond to the same proper value will then be determined only up to a linear transformation, so one can always choose those proper solutions in such a fashion that they will be orthogonal.

The case of a continuous spectrum of proper values will give rise to the same remarks as in the wave mechanics of a single function Ψ. Moreover, in a general fashion, the

parallelism between the two theories is complete here. Nevertheless, one must note that there is a difference: In the formulas that involve integration over space, such as the normalization and orthogonality formulas (3) and (4), one must sum over the index k in Dirac's theory. The presence of that summation will seem quite natural once we have developed the synthetic viewpoint, in which one considers the index k to be a sort of discontinuous supplementary variable (cf., Chapter XVI).

2. Matrices and first integrals in Dirac's theory. – In the wave mechanics of a single function Ψ, one makes any linear, Hermitian operator A correspond to the matrix whose elements are defined by the formula:

$$A_{mn} = \int_D \Psi_m^* \, A(\Psi_n) \, d\tau , \tag{14}$$

in which $d\tau$ is the volume element for the domain of existence D for the functions Ψ_n . One then says that the operator A is a *first integral* for the problem in question if all of the A_{mn} are independent of time, and we have seen that upon setting:

$$L = H - \frac{h}{2\pi i} \cdot \frac{\partial}{\partial t} ,$$

the necessary and sufficient for A to be a first integral will be that:

$$LA - AL = 0. \tag{15}$$

How should we translate these definitions into Dirac's theory?

In order to define the matrix elements, we take into account the remark that was made at the end of the last paragraph – i.e., that we suppose that a summation over the index of the Dirac functions Ψ enters into the integration that occurs in the old definition (14). The matrix elements that correspond to a linear, Hermitian operator A ([1]) will then be defined by the relation:

$$A_{mn} = \int_D \sum_{k=1}^{4} \Psi_{k,m}^* \, A(\Psi_{k,n}) \, d\tau . \tag{16}$$

We will always say that A is a first integral if all of the A_{mn} are independent of time. In order to find the condition that expresses the idea that A is a first integral, we write the symbolic Dirac equation in the condensed form:

$$L \, (\Psi) = 0, \tag{17}$$

with

([1]) Naturally, in Dirac's theory, an operator can operate upon the index k, as well as on the coordinates.

$$L = \left(P_4 + \sum_{i=1}^{3} \alpha_i P_i + \alpha_4 m_0 c \right), \tag{18}$$

and we remark that upon setting:

$$H = -\left[eV + \sum_{i=1}^{3} c\, \alpha_i P_i + \alpha_4\, m_0\, c^2 \right], \tag{19}$$

one can write:

$$L = \frac{1}{c}\left(-H + \frac{h}{2\pi i}\frac{\partial}{\partial t} \right), \qquad \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = H\,(\Psi). \tag{20}$$

The operator H that is defined by (19) is the Hamiltonian operator in Dirac's theory. The operator H is Hermitian; i.e., it satisfies the condition:

$$\int_D \sum_{k=1}^{4} \Psi^*_{k,m}\, H\,(\Psi_{k,n})\, d\tau = \int_D \sum_{k=1}^{4} \Psi_{k,n}\, H^*\,(\Psi^*_{k,m})\, d\tau \tag{21}$$

for all values of m and n. Equation (21) is the obvious generalization of the Hermiticity condition in the mechanics of one Ψ; moreover, it expresses simply that one has $H_{mn} = H^*_{mn}$, from the definition (16). Formula (21) is proved easily upon taking into account the Hermiticity of the α_i and the property of the $\Psi_{k,n}$ that they must be zero at the limits of D ([1]).

In order to express the idea that A is a first integral, we write:

$$\frac{\partial A_{mn}}{\partial t} = \int_D \sum_{k=1}^{4} \left[\Psi^*_{k,m} A\left(\frac{\partial \Psi_{k,n}}{\partial t} \right) + \Psi^*_{k,m}\frac{\partial A}{\partial t}(\Psi_{k,n}) + \frac{\partial \Psi^*_{k,m}}{\partial t} A(\Psi_{k,n}) \right] d\tau = 0. \tag{22}$$

([1]) In order to carry out that proof, one replaces the operator H in (21) with its expression (19), and one will verify that the relation (21) is valid for each term. That verification is immediate for the terms that contain V. For the terms that contain A_x, A_y, A_z, one will have, for example:

$$\sum_{k=1}^{4} e A_x \sum_{j=1}^{4} (\alpha_1^*)_{kj} \Psi_{k,n} = \sum_{j=1}^{4} e A_x \Psi^*_{j,m} \sum_{k=1}^{4} (\alpha_1^*)_{j,k} \Psi_{k,n} = \sum_{j=1}^{4} \Psi^*_{j,m} \alpha_1 \Psi_{j,n},$$

due to the Hermiticity of the α_i, and the relation (21) will thus be proved for those terms. What will remain are terms like:

$$\sum_{k=1}^{4} \frac{h}{2\pi i}\frac{\partial}{\partial x} \alpha_1^* \Psi^*_{k,m} \cdot \Psi_{k,n}.$$

One next has, due to the Hermiticity of the α_i, as always:

$$\frac{h}{2\pi i} \sum_{k,j=1}^{4} \frac{\partial}{\partial x}\left[\alpha_1^* \Psi^*_{k,m} \right] \Psi_{k,n} = \frac{h}{2\pi i} \sum_{j=1}^{4} \frac{\partial \Psi^*_{k,m}}{\partial x} \sum_{k=1}^{4} (\alpha_1)_{j,k} \Psi_{k,n},$$

and then upon integrating parts, one will have:

$$\frac{h}{2\pi i}\int_D \sum_{j=1}^{4} \frac{\partial \Psi^*_{j,m}}{\partial x} \sum_{k=1}^{4}(\alpha_1)_{j,k} \Psi_{k,n}\, d\tau = -\frac{h}{2\pi i}\int_D \sum_{j=1}^{4} \Psi^*_{j,m} \sum_{k=1}^{4}(\alpha_1)_{j,k} \frac{\partial \Psi_{k,n}}{\partial x}\, d\tau = \int_D \sum_{k=1}^{4} \Psi^*_{j,m}\left(-\frac{h}{2\pi i}\frac{\partial}{\partial x} \right)\alpha_1 \Psi_{k,n}\, d\tau,$$

because the $\Psi_{k,n}$ are zero on the boundary of D. Q. E. D.

Next, since we have $L(\Psi_{k,n}) = 0$ and $L^*(\Psi^*_{k,m}) = 0$, we replace the derivatives $\dfrac{\partial \Psi_{k,n}}{\partial t}$

and $\dfrac{\partial \Psi^*_{k,m}}{\partial t}$ with $\dfrac{2\pi i}{h} H(\Psi_{k,n})$ and $-\dfrac{2\pi i}{h} H^*(\Psi^*_{k,m})$, respectively, which will give us:

$$\frac{\partial A_{mn}}{\partial t} = \int_D \sum_{k=1}^4 \left[\frac{2\pi i}{h} \Psi^*_{k,m} A \Psi_{k,n} + \Psi^*_{k,m} \frac{\partial A}{\partial t}(\Psi_{k,n}) - \frac{2\pi i}{h} \Psi^*_{k,m} A(\Psi_{k,n}) \right] d\tau. \tag{23}$$

One can transform that expression with the aid of the formula:

$$\int_D \sum_{k=1}^4 H^*(\Psi^*_{k,m}) A(\Psi_{k,n}) \, d\tau = \int_D \sum_{k=1}^4 \Psi^*_{k,m} HA(\Psi_{k,n}) \, d\tau, \tag{24}$$

which is proved like (21). Upon substituting (24) in (23), one will find that:

$$\frac{\partial A_{mn}}{\partial t} = \int_D \sum_{k=1}^4 \Psi^*_{k,m} \left[\frac{\partial A}{\partial t} + \frac{2\pi i}{h}(AH - HA) \right] \Psi_{k,n} \, d\tau. \tag{25}$$

One then concludes from (22) that the necessary and sufficient condition for A to be a first integral is that:

$$\frac{\partial A}{\partial t} + \frac{2\pi i}{h}(AH - HA) = 0, \tag{26}$$

because the $\Psi_{k,n}$ form a complete system.

One can further write the condition (26) in a different way upon remarking that one has:

$$\frac{\partial}{\partial t} A(f) = \frac{\partial A}{\partial t}(f) + A\left(\frac{\partial f}{\partial t} \right) \tag{27}$$

for an arbitrary function f, so one will have, symbolically:

$$\frac{\partial A}{\partial t} = \frac{\partial}{\partial t} \cdot A - A \cdot \frac{\partial}{\partial t}. \tag{28}$$

Therefore, upon replacing the operator $\partial A / \partial t$ with its value (28) in (26), one will have:

$$L A - A L = 0, \tag{29}$$

by virtue of (26).

The conditions (26) and (29) have the same form as the analogous conditions in the mechanics of one Ψ, but with a different definition for the operators L and H.

We remark that equation (1) can be written in the form:

$$H(\Psi_k) = W \Psi_k \tag{30}$$

upon introducing the definition (19) of H.

The proper values W_k that were defined in paragraph 1 will then indeed be the proper values of the Hamiltonian operator – i.e., the proper values of energy.

3. Examples of first integrals. Proper angular momentum of the electron. – We shall first examine the cases in which the Hamiltonian operator that corresponds to energy is a first integral. For that to be true, it is necessary that one must have:

$$LH - HL = \frac{1}{c}\left(-H + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)H - H\frac{1}{c}\left(-H + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right) \equiv 0, \qquad (31)$$

or, more simply, that:

$$\frac{\partial}{\partial t}H - H \cdot \frac{\partial}{\partial t} = \frac{\partial H}{\partial t} \equiv 0. \qquad (32)$$

The necessary and sufficient condition for the energy to be a first integral is then always that the external field (which is defined by the four functions V, A_x, A_y, A_z, here) must be independent of time. That is the theorem of conservation of energy in Dirac's mechanics.

One also easily sees that if the potential vector is zero, and if the scalar potential does not depend upon one of the coordinates – say, x – then the corresponding component of the quantity of motion $\left(p_x = -\frac{h}{2\pi i}\frac{\partial}{\partial x}\right)$ will be a first integral.

The study of the theorem of the conservation of the moment of the quantity of motion is much more interesting in Dirac's theory. Indeed, we have seen that in the wave mechanics of a single function Ψ, the angular momentum around the z axis, which corresponds to the operator:

$$M_z = x\,p_y - y\,p_x = \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right), \qquad (33)$$

will be a first integral when the force field possesses cylindrical symmetry around the z axis. Now, we shall see that this result does not persist in the mechanics of four functions, and we shall now look for what should replace it.

We first verify that in a field that is derived from a scalar potential with cylindrical symmetry $V(\rho, z)$ with:

$$\rho = \sqrt{x^2 + y^2}\,,$$

M_z is not a first integral in Dirac's mechanics. In order to do that, we must show that M_z does not commute with L. It is obvious from the outset that M_z commutes with $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$,

$\alpha_3 \dfrac{h}{2\pi i} \dfrac{\partial}{\partial z}$, and $\alpha_4 \, m_0 c$. It also commutes with the term $\dfrac{e}{c} V$, because one will have (in operators):

$$V\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right) - \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right)V \;=\; x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\,, \tag{34}$$

which will be zero, because by hypothesis, V will depend upon only:

$$\rho = \sqrt{x^2 + y^2}\,.$$

However, by contrast, M_z does not commute with the terms in $-\alpha_1 \dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$ and $-\alpha_2 \dfrac{h}{2\pi i}\dfrac{\partial}{\partial y}$, because one will have:

$$\left(-\alpha_1 \frac{h}{2\pi i}\frac{\partial}{\partial x} - \alpha_2 \frac{h}{2\pi i}\frac{\partial}{\partial y} \right)\frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right) -$$

$$\frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right)\left(-\alpha_1 \frac{h}{2\pi i}\frac{\partial}{\partial x} - \alpha_2 \frac{h}{2\pi i}\frac{\partial}{\partial y} \right) = \frac{h^2}{4\pi^2}\left(\alpha_2 \frac{\partial}{\partial x} - \alpha_1 \frac{\partial}{\partial y} \right). \tag{35}$$

One will have, in total:

$$L\,M_z - M_z\,L = \frac{h^2}{4\pi^2}\left(\alpha_2 \frac{\partial}{\partial x} - \alpha_1 \frac{\partial}{\partial y} \right), \tag{36}$$

and M_z will not be a first integral.

Consider the operator:

$$N_z = M_z - \alpha_1\,\alpha_2\,\frac{h}{4\pi i} \tag{37}$$

then.

That operator is Hermitian, because the product of the two matrices α_1 and α_2, which are Hermitian and anti-commute, is anti-Hermitian, and the quotient by i is Hermitian. We shall show that N_z is a first integral in the case considered. In order to do that, we form the difference:

$$L\,\alpha_1\,\alpha_2\,\frac{h}{4\pi i} - \alpha_1\,\alpha_2\,\frac{h}{4\pi i}\,L\,.$$

The terms in $\partial/\partial t$, V, and $\alpha_4\,m_0\,c$, and L obviously commute with $\alpha_1\,\alpha_2$; the same thing will be true for the term $\alpha_3 \dfrac{h}{2\pi i}\dfrac{\partial}{\partial z}$, by virtue of the properties of α_1. What will then remain is:

$$L \, \alpha_1 \, \alpha_2 \frac{h}{4\pi i} - \alpha_1 \, \alpha_2 \frac{h}{4\pi i} L$$

$$= \frac{h^2}{8\pi^2} \left[\alpha_1 \alpha_1 \alpha_2 \frac{\partial}{\partial x} + \alpha_2 \alpha_1 \alpha_2 \frac{\partial}{\partial y} - \alpha_1 \alpha_2 \alpha_1 \frac{\partial}{\partial x} - \alpha_1 \alpha_2 \alpha_2 \frac{\partial}{\partial y} \right]. \tag{38}$$

Since one has $\alpha_1 \, \alpha_1 = \alpha_2 \, \alpha_2 = 1$, $\alpha_2 \, \alpha_1 \, \alpha_2 = - \, \alpha_1$ and $\alpha_1 \, \alpha_2 \, \alpha_1 = - \, \alpha_2$, one will find that:

$$L \, \alpha_1 \, \alpha_2 \frac{h}{4\pi i} - \alpha_1 \, \alpha_2 \frac{h}{4\pi i} L = \frac{h^2}{4\pi^2} \left(\alpha_2 \frac{\partial}{\partial x} - \alpha_1 \frac{\partial}{\partial y} \right) = L \, M_z - M_z \, L. \tag{39}$$

One then concludes the relation:

$$L \, M_z - M_z \, L \equiv 0 \tag{40}$$

from that, and one will see that N_z is a first integral.

One can present an argument that is analogous to the preceding one for the components with the indices x and y by permuting the roles of the axes. One will then see that in Dirac's theory, the *total* angular momentum of the electron is a vector with components:

$$N_x = M_x + S_x , \qquad N_y = M_y + S_y , \qquad N_z = M_z + S_z , \tag{41}$$

in which the operators that correspond to M_x, etc. are:

$$M_x = \frac{h}{2\pi i} \left(z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z} \right), \quad M_y = \frac{h}{2\pi i} \left(x \frac{\partial}{\partial z} - z \frac{\partial}{\partial x} \right), \quad M_z = \frac{h}{2\pi i} \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right)$$

$$\tag{42}$$

$$S_x = - \, \alpha_2 \, \alpha_3 \frac{h}{4\pi i} , \quad S_y = - \, \alpha_3 \, \alpha_1 \frac{h}{4\pi i} , \quad S_x = - \, \alpha_1 \, \alpha_2 \frac{h}{4\pi i} .$$

M_x, M_y, M_z are the components of the "orbital" angular momentum of the electron, and one will be led quite naturally to consider S_x, S_y, S_z to be the components of the proper angular momentum of the electron – i.e., the "spin."

4. Explicit calculation of N_z . Sign of the proper mass for the wave Ψ_k . – It is very interesting to calculate explicitly the operator N_z that corresponds to the privileged z axis and to see the results that one will get upon successively applying it to the four Ψ_k.

Upon starting with the values that we know for α_1 and α_2, we easily calculate the matrix product with the aid of the rules for the multiplication of matrices, and we will find that:

$$\alpha_1 \alpha_2 = \begin{vmatrix} -i & 0 & 0 & 0 \\ 0 & +i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & +i \end{vmatrix}. \tag{43}$$

One will easily infer from this that:

$$N_z(\Psi_1) = M_z(\Psi_1) + \frac{h}{4\pi}\Psi_1,$$

$$N_z(\Psi_2) = M_z(\Psi_2) - \frac{h}{4\pi}\Psi_2,$$

$$\tag{44}$$

$$N_z(\Psi_3) = M_z(\Psi_3) + \frac{h}{4\pi}\Psi_3,$$

$$N_z(\Psi_4) = M_z(\Psi_4) - \frac{h}{4\pi}\Psi_4.$$

We then see from formulas (44) that the proper angular momentum will be $+\dfrac{h}{4\pi}$ for the wave functions with odd indices, and $-\dfrac{h}{4\pi}$ for the ones with even indices.

We now write the mean value of N_z. We have:

$$\overline{N_1} = \iiint \sum_{k=1}^{4} \Psi_k^* N_z(\Psi_k)\, d\tau = \overline{M_z} + \frac{h}{4\pi}\iiint [\Psi_1^*\Psi_1 - \Psi_2^*\Psi_2 + \Psi_3^*\Psi_3 - \Psi_4^*\Psi_4]\, d\tau. \tag{45}$$

One must interpret that formula by saying: The proper angular momentum of the electron along the z axis can take on only the two values $\dfrac{h}{4\pi}$ and $-\dfrac{h}{4\pi}$, so the probability under the first hypothesis will be:

$$\iiint (\Psi_1^*\Psi_1 + \Psi_3^*\Psi_3)\, d\tau,$$

and the probability under the second hypothesis will be:

$$\iiint [\Psi_2^*\Psi_2 + \Psi_4^*\Psi_4]\, d\tau.$$

One then compares formula (45) with formula (16) in Chapter XIV, which gives the mean magnetic moment of the electron along oz:

$$\overline{M_z} = \iiint (\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4)\, d\tau \,. \tag{46}$$

We see that one can make the following correspondences:

the wave Ψ_3 :   a magnetic moment $-\dfrac{eh}{4\pi m_0 c}$

and an angular momentum $+\dfrac{h}{4\pi}$,

the wave Ψ_4 :   a magnetic moment $+\dfrac{eh}{4\pi m_0 c}$

and an angular momentum $-\dfrac{h}{4\pi}$.

For these two waves (which are dominant in the Newtonian approximation, as we have seen), the ratio of the two moments will be equal to $-\dfrac{e}{m_0 c}$, as one would expect from the double magnetism of the rotating electron ([1]).

However, we will obtain a surprising result for the wave Ψ_1 and Ψ_2. In fact, from (45) and (46), one can make the following correspondences:

the wave Ψ_1 :   a magnetic moment $+\dfrac{eh}{4\pi m_0 c}$

and an angular momentum $+\dfrac{h}{4\pi}$,

the wave Ψ_2 :   a magnetic moment $-\dfrac{eh}{4\pi m_0 c}$

and an angular momentum $-\dfrac{h}{4\pi}$,

so the ratio of the two moments will be $+\dfrac{e}{m_0 c}$, which differs by a sign from the expected value. Where does that anomaly come from? In order to understand that, we first remark that everything will take place as if the proper mass of the electron were $- m_0$, instead of m_0, as far as the wave Ψ_1 and Ψ_2 are concerned. One will certainly see this in the Dirac equations themselves:

$$\left(P_4 + \sum_{i=1}^{4} \alpha_i P_i + \alpha_4 m_0 c \right) \Psi_4 = 0 \qquad (k = 1, 2, 3, 4), \tag{47}$$

([1])   The ratio $- e / m_0\, c$ is indeed twice the normal ratio that was given by formula (7) in Chapter IV.

when one retains the explicit form for the matrix α_4 :

$$\alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}. \tag{48}$$

The term m_0 has a different sign in the first two equations (47), on the one hand, and the last two, on the other.

One can express that fact by saying that in Dirac's theory, the proper mass m_0 of the old mechanics is replaced with the operator $-\alpha_4\, m_0$. Since one has:

$$-\alpha_4\, m_0\, \Psi_1 = -m_0\, \Psi_1 , \qquad -\alpha_4\, m_0\, c\, \Psi_2 = -m_0\, \Psi_2 ,$$

$$-\alpha_4\, m_0\, c\Psi_3 = m_0\, \Psi_3 , \qquad -\alpha_4\, m_0\, c\, \Psi_4 = m_0\, \Psi_4 , \tag{49}$$

it is obvious that the waves Ψ_3 and Ψ_4 will correspond to the mass m_0 , while the waves Ψ_1 and Ψ_2 will correspond to the mass $-m_0$. If the reader would now like to review the argument that preceded formula (32) in Chapter X then he will comprehend its importance better in light of what was just said.

———————

CHAPTER XVI

SYSTEMATIC SUMMARY OF THE RESULTS OBTAINED
___

1. Considering the index of the function Ψ_k to be a variable.

Up to now, we have presented Dirac's theory by saying that it involves the existence of four wave functions Ψ_1, Ψ_2, Ψ_3, Ψ_4 of the four variables x, y, z, t. We can take a different viewpoint and say that there is *just one* function Ψ that depends upon the four continuous variables x, y, z, t that are capable of taking on all of the real values from $-\infty$ to $+\infty$, and a fifth variable ζ – namely, "spin" – that can take on the four values 1, 2, 3, 4. That amounts to considering the index k of the functions Ψ_k to be a discontinuous variable with four possible values.

The operators α_i operate on the discontinuous variable ζ, while the operators like p_x, for example, act upon the continuous variables. It is legitimate to consider operators that operate on all four variables ([1]) x, y, z, t, such as the Dirac Hamiltonian H that was defined by formula (19) of the last chapter.

Now write the Dirac equation in the form:

$$\left(P_4 + \sum_{i=1}^{3} \alpha_i P_i + \alpha_4 m_0 c \right) \Psi\,(x,\,y,\,z,\,t,\,\zeta) = 0. \qquad (1)$$

Up to now, we have regarded that equation as a symbolic equation that summarizes four distinct equations. We can now consider it to be a single equation of propagation through the space of four variables x, y, z, ζ in the course of time.

The integrations that one performs over the space of three dimensions x, y, z in wave mechanics "without spin" must be performed over the space of x, y, z, ζ, here. That will explain the deep reason for the following fact, which was pointed out before: All of the formulas from wave mechanics in which an integration over space occurs must be completed with a summation over the index k from 1 to 4 in Dirac's theory.

In effect, that summation $\sum_{k=1}^{4}$ corresponds to a sort of integration over the discontinuous variable ζ, and it will often be convenient for us to represent it by an integral $\int \ldots d\zeta$. Thus, for example, the normalization condition in wave mechanics without spin:

$$\int\int\int_{-\infty}^{+\infty} \Psi^*(x,\,y,\,z,\,t) \cdot \Psi(x,\,y,\,z,\,t)\,dx\,dy\,dz = 1 \qquad (2)$$

will become:

___

([1]) Although Dirac's theory is relativistic, in a sense, time plays a different role from the other four variables x, y, z, ζ in it; we shall return to that point later on.

$$\int\int\limits_{-\infty}^{+\infty}\int \sum_{k=1}^{4} \Psi_k^* \cdot \Psi_k \, dx \, dy \, dz$$

$$= \iiint\int \ \int\int\limits_{-\infty}^{+\infty}\int \ \Psi^*(x, y, z, t, \zeta) \cdot \Psi(x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta = 1 \qquad (3)$$

in Dirac's theory.

Similarly, the new viewpoint permits one to easily see how one must write the series developments in proper functions in Dirac's theory. Indeed, let $f(x, y, z, t, \zeta)$ be a function of the five variables x, y, z, t, ζ. It is equivalent to four functions f_1, f_2, f_3, f_4 of the four continuous variables x, y, z, t. Suppose that we know the system of proper functions of the Hamiltonian operator. We have already agreed to denote a proper function of that system by the set of four functions $\Psi_{1,m}$, $\Psi_{2,m}$, $\Psi_{3,m}$, $\Psi_{4,m}$, in which the second index characterizes the corresponding proper value. From our new viewpoint, we must represent the set of four functions $\Psi_{k,m}$ by the single function $\Psi_m(x, y, z, t, \zeta)$.

The development of the function $f(x, y, z, t, \zeta)$ in the complete system of $\Psi_m(x, y, z, t, \zeta)$ is then written quite naturally as:

$$f(x, y, z, t, \zeta) = \sum_m c_m \, \Psi_m(x, y, z, t, \zeta), \qquad (4)$$

which is equivalent to the four relations:

$$f_k(x, y, z, t) = \sum_m c_m \, \Psi_{k,m}(x, y, z, t) \qquad (k = 1, 2, 3, 4). \qquad (5)$$

One then sees that each of the four components f_k are developed in proper functions $\Psi_{k,m}$ with *the same index k*, and in addition the coefficients of the developments are the same for the four components. That proposition, which might not seem to be obvious when one considers k to be an index, seems, in the contrary, quite natural when one argues as we just did by replacing the index k with the discontinuous variable ζ.

2. Statement of the general principles in Dirac's theory. – The introduction of the discontinuous variable ζ permits one to immediately find how the general principles of wave mechanics transpose into Dirac's theory.

First of all, we assume that any observable physical quantity that is attached to the electron corresponds to an operator A that can generally operate on all four variables x, y, z, ζ. That operator must always be Hermitian in the space of x, y, z, ζ; i.e., one must have:

$$\iiint\int f^*(x, y, z, t, \zeta) A g(x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta$$

$$= \iiint\int f(x, y, z, t, \zeta) A^* g^*(x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta. \qquad (6)$$

The proper values of the operator A are defined to be the values of the constant a for which the equation:

$$A \, \varphi(x, y, z, t, \zeta) = a \, \varphi(x, y, z, t, \zeta) \qquad (7)$$

admits at least one everywhere-finite, continuous, and uniform solution in x, y, z for each of the four values of ζ. These proper values, which are generally functions of time, are real, and two proper functions φ_m and φ_n that correspond to two distinct proper values a_m and a_n will be orthogonal in the space x, y, z, ζ; i.e., one will have:

$$\iiint \varphi_m^*(x, y, z, t, \zeta) \, \varphi_n(x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta = 0. \qquad (8)$$

We shall prove these results by the same method that we employed in Chapter V, paragraph 4, by always adding a summation over ζ to the integrations over x, y, z. Moreover, we shall always suppose that the proper functions φ_n have been normalized by the condition:

$$\iiint \varphi_n^*(x, y, z, t, \zeta) \, \varphi_n(x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta = 1. \qquad (9)$$

Having said that, it is easy to state the fundamental principles of the new mechanics in the context of Dirac's theory.

The first of these principles can be stated without changing anything by saying that a measurement of the quantity A that is performed at the instant t will necessarily yield one of the proper values of the operator A at that instant as a result.

In order to state the second principle, we suppose, first of all, that the operator A is a *complete* operator; i.e., that it involves all four variables x, y, z, ζ. Furthermore, we also suppose that it is *nondegenerate;* i.e., that it has no multiple proper values. Therefore, let $\Psi(x, y, z, t, \zeta)$ be the wave function of an electron. That wave function can be developed in proper functions of the operator A in the form:

$$\Psi(x, y, z, t, \zeta) = \sum_m c_m \, \varphi_n(x, y, z, t, \zeta), \qquad (10)$$

in which the c_m are complex constants that are generally functions of time. The second principle then asserts that $| c_m(t) |^2$ is the probability that a measurement of the quantity A will yield the proper value a_m that corresponds to φ_m.

If A possesses multiple proper values then the same a_m will correspond to several φ_m. The probability of getting the value a_m for the quantity A will then be equal to the sum of the squares of the moduli of the coefficients that relate to these φ_m in the development of Ψ.

If the operator A is incomplete – i.e., it involves only some of the variables x, y, z, ζ – then the corresponding φ_m will depend upon only those variables and coefficients c_m in the development of Ψ in φ_m that depend upon variables that do not figure in A. In that case, if one is to obtain the probability of the proper value a_m then it will be necessary to integrate the quantity $| c_m(t) |^2$ over the entire domain of those variables x, y, z, ζ upon which c_m depends. For example, if A is an operator that acts only upon the variable ζ,

like one of the α_i, then those proper functions will have the form $\varphi_m(\zeta)$, and one will have:

$$\Psi(x, y, z, t, z) = \sum_m c_m(x, y, z, t)\, \varphi_m(\zeta). \tag{11}$$

The probability that the quantity A has the value a_m will then be given by:

$$\int\int\int_{-\infty}^{+\infty} |c_m(x, y, z, t)|^2\, dx\, dy\, dz.$$

We will see an application of that result in the following paragraph.

Having assumed those principles, one will see that the mean value of the quantity A is indeed, equal to:

$$A = \int\int\int\int \Psi^* A(\Psi)\, dx\, dy\, dz\, d\zeta = \int\int\int_{-\infty}^{+\infty} \Psi_k^* A(\Psi_k)\, d\tau, \tag{12}$$

as we assumed, because if one replaces Ψ and Ψ^* in that expression with their developments in proper functions of A then one will easily see that it is equal to the sum of the products of each proper value with its probability. The quantity $\sum_{k=1}^{4} \Psi_k^* A(\Psi_k)$ can then be called the "mean-value density" for the quantity A, but as we pointed out in Chapter VI, paragraph 3, such a density cannot be considered to have the same physical sense as the densities in classical theories. Nevertheless, we shall soon see that the densities of that type that correspond to the operators α_i and to Hermitian operators that are composed of products of the α_i will be real and possess a tensorial character that permits us to get closer to certain quantities of classical physics.

Finally, recall that the element ij of the matrix that corresponds to an operator A must be defined in Dirac's mechanics by the formula:

$$A_{ij} = \int\int\int\int \Psi_i^*(x, y, z, t, \zeta)\, A\left(\Psi_i(x, y, z, t, \zeta)\right) dx\, dy\, dz\, d\zeta$$

$$= \int\int\int_{-\infty}^{+\infty} \sum_{k=1}^{4} \Psi_{k,i}^* A(\Psi_{k,j})\, d\tau, \tag{13}$$

and is equal to the coefficient of $\Psi_i(x, y, z, t, \zeta)$ in the development of the function $A(\Psi_i)$ in proper functions of the Hamiltonian operator.

In the foregoing, we implicitly supposed that the spectrum of the operator A was purely discontinuous. What we said in Chapters V and VI about continuous spectra will permit the reader to see effortlessly how the formulas that were written above will be modified when there exist proper values that define a continuous sequence.

3. Example: proper values of the operator $\mathfrak{M}_z = \dfrac{e\,h}{4\pi\,m_0 c}\,i\,\alpha_1\,\alpha_2\,\alpha_4$. – As an example of an operator A that acts upon only the spin variable, we take the operator $\mathfrak{M}_z = \dfrac{e\,h}{4\pi\,m_0 c}\,i\,\alpha_1\,\alpha_2\,\alpha_4$, which we have made to correspond to the z component of the proper magnetic moment of the electron (cf., Chapter IX, at the end).

In order to abbreviate the writing, we set:

$$ B = \frac{e\,h}{4\pi\,m_0 c} = \text{Bohr magneton.} \tag{14} $$

From (7), the equation that defines the proper values and proper functions of the operator $\mathfrak{M}_z$ is:

$$ B \sum_{k=1}^{4} i\,(\alpha_1\,\alpha_2\,\alpha_4)_{jk}\,\varphi_j = a\,\varphi_j \qquad (j = 1,\,2,\,3,\,4). \tag{15} $$

Upon taking into account the values of the elements of the matrices α_1, α_2, α_4 , one will find that for each of the four values of the variable ζ, one will have:

$$ B\,\varphi\,(1) = +\,\varphi\,(1), \quad B\,\varphi\,(2) = -\,\varphi\,(2), \quad B\,\varphi\,(3) = -\,\varphi\,(3), \quad B\,\varphi\,(4) = +\,\varphi\,(4). \tag{16} $$

These equations will admit a solution $\varphi\,(\zeta)$ that is not identically zero only if one has $a = \pm\,B$. The operator $\mathfrak{M}_z$ will then have only the values $\pm\,B$ for proper values, as we expected.

For the proper value $a = +\,B$, there exist two *independent* proper functions that we call:

$$ \varphi_+^{(1)}(\zeta) \quad \text{and} \quad \varphi_+^{(2)}(\zeta). $$

These proper functions are defined by their values for the four possible values of the variable ζ, namely:

$$ \varphi_+^{(1)}(1) = 1, \qquad \varphi_+^{(1)}(2) = 0, \qquad \varphi_+^{(1)}(3) = 0, \qquad \varphi_+^{(1)}(4) = 0, $$
$$ \varphi_+^{(2)}(1) = 0, \qquad \varphi_+^{(2)}(2) = 0, \qquad \varphi_+^{(2)}(3) = 0, \qquad \varphi_+^{(2)}(4) = 1. \tag{17} $$

These functions are normalized and orthogonal because one has:

$$ \int \left| \Psi_+^{(1)} \right|^2 d\zeta = 1, \qquad \int \left| \Psi_+^{(2)} \right|^2 d\zeta = 1, \qquad \int \Psi_+^{(1)*}\,\Psi_+^{(2)}\,d\zeta = 0. \tag{18} $$

The proper value $+\,B$ is then a double one.

Similarly, the proper value $a = -B$ is double, because it corresponds to two independent proper functions that are normalized and orthogonal. $\varphi_-^{(1)}(\zeta)$ and $\varphi_-^{(2)}(\zeta)$ are defined by:

$$\varphi_-^{(1)}(1) = 0, \qquad \varphi_-^{(1)}(2) = 1, \qquad \varphi_-^{(1)}(3) = 0, \qquad \varphi_-^{(1)}(4) = 0,$$

$$\varphi_-^{(2)}(1) = 0, \qquad \varphi_-^{(2)}(2) = 0, \qquad \varphi_-^{(2)}(3) = 1, \qquad \varphi_-^{(2)}(4) = 0. \tag{19}$$

One verifies that the two functions φ_- are orthogonal to the functions φ_+.

Therefore, let $\Psi(x, y, z, t, \zeta)$ be the wave function of an electron. By virtue of the first general principle, the possible values of the z component of its proper magnetic moment are the proper values of the operator $\mathfrak{M}_z$, namely, $\pm B$; i.e., ± 1 times the Bohr magneton. In order to find the probabilities for these two respective hypotheses, we must write the development as:

$$\Psi(x, y, z, t, \zeta) = c_+^{(1)}(x, y, z, t)\,\varphi_+^{(1)}(x, y, z, t) + c_+^{(2)}(x, y, z, t)\,\varphi_+^{(2)}(x, y, z, t)$$
$$+ c_-^{(1)}(x, y, z, t)\,\varphi_-^{(1)}(x, y, z, t) + c_-^{(2)}(x, y, z, t)\,\varphi_-^{(2)}(x, y, z, t). \tag{20}$$

From the second general principle, the probability of the vector $+ B$ is:

$$\int\!\!\int\!\!\int_{-\infty}^{+\infty} \left[\left|c_+^{(1)}\right|^2 + \left|c_+^{(2)}\right|^2\right] d\tau,$$

and similarly that of the value $- B$ is:

$$\int\!\!\int\!\!\int_{-\infty}^{+\infty} \left[\left|c_-^{(1)}\right|^2 + \left|c_-^{(2)}\right|^2\right] d\tau.$$

Now, upon successively substituting the values $\zeta = 1, 2, 3, 4$ in formula (20), one will easily find that:

$$c_+^{(1)} = \Psi_1(x, y, z, t), \qquad c_+^{(2)} = \Psi_4(x, y, z, t),$$

$$c_-^{(1)} = \Psi_2(x, y, z, t), \qquad c_-^{(2)} = \Psi_3(x, y, z, t). \tag{21}$$

The probability of the value $+ B$ will then be:

$$\int\!\!\int\!\!\int_{-\infty}^{+\infty} \left[\left|\Psi_1\right|^2 + \left|\Psi_4\right|^2\right] d\tau,$$

and that of the value $- B$ will be:

$$\int\int\limits_{-\infty}^{+\infty}\int \left[|\Psi_2|^2 + |\Psi_3|^2 \right] d\tau.$$

That is indeed the conclusion to which we arrived before.

The simple example that we just treated shows how one can obtain the values of the proper functions of the operators that act upon only the ζ. In particular, those considerations are applicable to the operators that we shall speak of at the beginning of the following paragraph, which are operators whose double proper values have the values ± 1.

4. The 16 fundamental operators of Dirac's theory. Corresponding quantities and densities. – With the aid of the operators α_i and the matrix 1 (with four rows and four columns), one can form the following table that contains sixteen Hermitian operators that act upon only the variable ζ:

$$
\begin{array}{c}
\alpha_4 \\
\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad 1 \\
i\,\alpha_2\,\alpha_3\,\alpha_4 \quad i\,\alpha_3\,\alpha_1\,\alpha_4 \quad i\,\alpha_1\,\alpha_2\,\alpha_4 \quad i\,\alpha_1\,\alpha_4 \quad i\,\alpha_2\,\alpha_4 \quad i\,\alpha_3\,\alpha_4 \\
i\,\alpha_2\,\alpha_3 \quad i\,\alpha_3\,\alpha_1 \quad i\,\alpha_1\,\alpha_2 \quad i\,\alpha_1\,\alpha_2\,\alpha_3 \\
\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4\,.
\end{array}
\qquad (22)
$$

In that table, one has multiplied the products of the α_i by i when they are anti-Hermitian, in such a fashion as to obtain Hermitian operators. Naturally, one can obtain further operators – such as $i\,\alpha_4\,\alpha_1$, for example – upon permuting the α_i, but each of the new operators will be equal or equal with the opposite sign to one of the operators that appear in the table.

Upon multiplying some of the operators in the table (22) by a suitable factor, we will find operators that correspond to the quantities that were studied before. Therefore, from formulas (7) and (8) of Chapter XII, the operators of the second row will correspond to the components of the mean electric current and the mean electric density when the first three are multiplied by ec, and the last one, by $-e$, resp.

From equations (9) of Chapter XIV, the operators of the third row will correspond to the coordinates of the proper magnetic moment and proper electric moment of the electron when they are multiplied by the Bohr magneton $\dfrac{eh}{4\pi m_0 c}$. Finally, from formulas (42) of Chapter XV, the first three operators of the fourth row will correspond to the three components S_x, S_y, S_z of the proper angular momentum when they are multiplied by $h/4\pi$. From what we said at the end of the last chapter, we shall also propose that the operator α_4 can indeed correspond to the proper mass when it is multiplied by $-m_0$.

With the 16 operators of the table (22), we can define 16 mean-value densities that have a tensorial character and are all real, and which we can write as follows:

$$\Omega_1 = \sum_{k=1}^{4} \Psi_k^* \, \alpha_4 \Psi_k \qquad\qquad (a)$$

$$i_1 = ec \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \Psi_k \qquad i_2 = ec \sum_{k=1}^{4} \Psi_k^* \, \alpha_2 \Psi_k$$

$$i_3 = ec \sum_{k=1}^{4} \Psi_k^* \, \alpha_3 \Psi_k \qquad i_4 = - ec \sum_{k=1}^{4} \Psi_k^* \, 1 \Psi_k \qquad\qquad (b)$$

$$\mu_{32} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_2 \alpha_3 \alpha_4 \Psi_k \qquad \mu_{14} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \alpha_4 \Psi_k \qquad \mu_{13} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_3 \alpha_1 \alpha_4 \Psi_k$$

$$\mu_{24} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_2 \alpha_4 \Psi_k \qquad \mu_{21} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \alpha_2 \alpha_4 \Psi_k \qquad \mu_{34} = \frac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_3 \alpha_4 \Psi_k \qquad (c) \quad (23)$$

$$\sigma_1 = \frac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_2 \alpha_3 \Psi_k \qquad \sigma_2 = \frac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_3 \alpha_1 \Psi_k$$

$$\sigma_3 = \frac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \alpha_2 \Psi_k \qquad \sigma_4 = \frac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \alpha_2 \alpha_3 \Psi_k \qquad\qquad (d)$$

$$\Omega_2 = \sum_{k=1}^{4} \Psi_k^* \, \alpha_1 \alpha_2 \alpha_3 \alpha_5 \Psi_k \qquad\qquad (e)$$

The quantities Ω_1 and Ω_2 are invariant under Lorentz transformations; we already know that. When envisioned as mean-value densities, they will be invariant densities. The physical significance (if it exists) of Ω_2 is still unknown; that of the quantity Ω_1 will be discussed later on.

The four quantities (b) form a space-time quadri-vector ([1]). We already know that it corresponds to the four components of the "world electric current" that is well-known in relativity. Upon integrating the component i_4 over all space, we will obtain an invariant quantity that is the total electric charge $-e$ of the electron, up to a factor of c.

The six quantities (c) are the six distinct components of a second-order anti-symmetric tensor. In fact, we already know of it: They are the magnetic moment and electric moment densities that were studied in Chapter XIV, paragraph 3.

The four quantities (d) transform like the components of a space-time vector. The first three of them σ_x, σ_y, σ_z are, as we know already, the mean-value densities of the proper angular momentum $\mathbf{S}$ (i.e., the spin). The temporal component σ_4 completes the space-time vector; its physical interpretation does not seem to be very clear.

([1])   In table (23), the components with an index 4 refer to the fourth space-time variable, namely, $x_4 = ct$. One will then have $i_4 = c\rho$.

5. Remarks on the vector $\boldsymbol{\sigma}$. – From the preceding, the space-time vector $\boldsymbol{\sigma}$ can be called the *spin vector.* By definition, the length $| \boldsymbol{\sigma} |$ of such a space-time vector is given by the formula:

$$| \sigma |^2 = \sigma_4{}^2 - \sigma_1{}^2 - \sigma_2{}^2 - \sigma_3{}^2. \tag{24}$$

If one does the calculation then the result will be:

$$| \sigma |^2 = -(\Omega_1{}^2 + \Omega_2{}^2)\frac{h^2}{16\pi^2}. \tag{25}$$

The spin is then a spacelike vector, as one says in relativity theory. That is clearly distinguished from the current vector, which is timelike, from formula (14) of Chapter XII.

Now recall the monochromatic plane wave with the wave function:

$$\Psi_1 = - \frac{Ap_z}{W/c + m_0 c} e^{\frac{2\pi i}{h}(Wt - p_z z)},$$

$$\Psi_2 = \frac{Bp_z}{W/c + m_0 c} e^{\frac{2\pi i}{h}(Wt - p_z z)},$$

$$\Psi_3 = A e^{\frac{2\pi i}{h}(Wt - p_z z)}, \tag{26}$$

$$\Psi_4 = B e^{\frac{2\pi i}{h}(Wt - p_z z)},$$

in which $W = c \sqrt{m_0^2 c^2 + p_z^2}$.

We carry out the explicit calculation of the components of $\boldsymbol{\sigma}$ in this particular case; we will find that:

$$\sigma_1 = \frac{h}{4\pi}(A^* B + AB^*)\left(1 - \frac{p_z^2}{(W/c + m_0 c)^2}\right),$$

$$\sigma_2 = \frac{h}{4\pi}i(A^* B - AB^*)\left(1 - \frac{p_z^2}{(W/c + m_0 c)^2}\right),$$

$$\sigma_3 = \frac{h}{4\pi}(AA^* - BB^*)\left(1 + \frac{p_z^2}{(W/c + m_0 c)^2}\right), \tag{27}$$

$$\sigma_4 = \frac{h}{4\pi}2(AA^* - BB^*)\frac{p_z}{W/c + m_0 c}.$$

Since the components of the "world electric current" quadri-vector for a plane wave are:

$$i_1 = j_x = 0, \qquad i_2 = j_y = 0, \qquad i_3 = j_z = -e\,\rho\,\frac{c^2\,p_z}{W}, \quad i_4 = c\,\delta = -e\,\rho\,c,$$

which results from the formulas of Chapter XII, para. 4, we will easily find that:

$$(\boldsymbol{i} \cdot \boldsymbol{\sigma}) = i_4\,\sigma_4 - i_1\,\sigma_1 - i_2\,\sigma_2 - i_3\,\sigma_3 = 0. \tag{28}$$

Therefore, the space-time scalar product of the vectors $\boldsymbol{i}$ and $\boldsymbol{\sigma}$ is zero. Hence, for the monochromatic plane wave, the two quadri-vectors "world current" and "spin" are orthogonal in space-time. (Of course, that is not to say that the corresponding spatial vectors are perpendicular.)

One can also write (28) in the form:

$$\sigma_4 = \sigma_3\,\frac{p_z c}{W} = \sigma_1\,v_1 + \sigma_2\,v_2 + \sigma_3\,v_3 . \tag{28 cont.}$$

The component σ_4 is thus interpreted here as the *spatial* scalar product of the "proper angular momentum density" and "velocity" vectors.

For a speed that is close to c, the ratio $\dfrac{p_z}{W/c + m_0 c}$ will be close to 1, and from (27), the components σ_x, σ_y will be almost zero. Upon confronting this result with what was said at the end of the last chapter, one will see that in Dirac's mechanics, a particle that is animated with a velocity that is close to c and is associated with a monochromatic plane wave will have its three spatial vectors $\mathbf{I}$, $\mathbf{J}$, and $\boldsymbol{\sigma}$ mutually perpendicular, and the last one will directed along the normal to the wave [1].

We conclude these remarks on spin by pointing out a general relation between the components of that vector and the invariant Ω_2 . That relation, which is due to Uhlenbeck and Laporte, is written:

$$\frac{1}{c}\frac{\partial\sigma_4}{\partial t} + \frac{\partial\sigma_1}{\partial x} + \frac{\partial\sigma_2}{\partial y} + \frac{\partial\sigma_3}{\partial z} = \sum_{i=1}^{4}\frac{\partial\sigma_i}{\partial x_i} = -m_0\,c\,\Omega_2 \tag{29}$$

with our notations.

---

[1] One might be surprised to see that the proper angular momentum $\mathbf{S}$ of the Dirac electron does not always coincide in direction with the magnetic moment $\mathfrak{M}$ and believe that this contradicts the relation:

$$\frac{\mathfrak{M}}{S} = -\frac{e}{m_0 c}, \tag{*}$$

by which, one expresses the double magnetism of the electron. However, must remark that the relation (*) is valid only for reference systems in which the electron is rest, because $\mathfrak{M}$ and $\mathbf{S}$ do not transform the same under a Lorentz transformation. Indeed, for a magnetic moment, it is the *densities* of the components that transform like the components 23, 32, and 12 of a tensor; on the contrary, for an angular momentum, it is the components themselves that transform in that way.

In the case of the plane wave, it is satisfied identically because all of the terms are zero separately. One proves the relation (28) easily by starting with the Dirac equation and its conjugate; its true physical significance remains unknown.

6. Remarks on the invariant Ω_1 and the operator $- m_0 \alpha_4$. – We have seen already that, in a certain sense, the physical quantity "proper mass" will correspond to the operator $- m_0 \alpha_4$ under the Dirac equations. If we assume that correspondence then the mean-value density that one deduces from it will be:

$$- m_0 \sum_{k=1}^{4} \Psi_k^* \alpha_4 \Psi_k \; = - m_0 \Omega_1 \; . \tag{30}$$

It is invariant. If one integrates that density over space then one will obtain the mean value:

$$\overline{m}_0 \; = - m_0 \int_{+\infty}^{-\infty}\!\!\int\!\!\int \sum_{k=1}^{4} \Psi_k^* \alpha_4 \Psi \, d\tau = - m_0 \int_{-\infty}^{+\infty}\!\!\int\!\!\int \Omega_1 \, d\tau \; , \tag{31}$$

or, upon making Ω_1 explicit:

$$\overline{m}_0 \; = m_0 \int_{-\infty}^{+\infty}\!\!\int\!\!\int (\Psi_3^* \Psi_3 + \Psi_4^* \Psi_4 - \Psi_1^* \Psi_1 - \Psi_2^* \Psi_2) \, d\tau \; , \tag{32}$$

which is a formula that indeed confirms the idea that the wave functions with indices 3 and 4 correspond to the proper mass $+ m_0$, and those with indices 1 and 2, to the proper mass $- m_0$.

The mean value (31) is not invariant. If one would like to deduce an invariant quantity from it then one must integrate over the fourth space-time variable $x = ct$. One will then obtain:

$$\int^t m_0 c \, dt \; = - m_0 \, c \int^t dt \int_{-\infty}^{+\infty}\!\!\int\!\!\int \Omega_1 \, d\tau \; . \tag{33}$$

In order to try to penetrate the physical significance of the formulas a little, we again recall the case of the monochromatic plane wave (26) and calculate the value that $\overline{m}_0$ will have in that case. We easily find from (30) that:

$$- m_0 \, \Omega_1 = m_0 \, (AA^* + BB^*) \left[1 - \frac{p_z^2}{(W/c + m_0 c)^2} \right] = m_0 \, (AA^* + BB^*) \, \frac{2 m_0 \, c^2}{W/c + m_0 c} \; . \tag{34}$$

In order to get $\bar{m}_0$, one must integrate the expression that this will give upon taking into account the normalization of the Ψ [1]:

$$\bar{m}_0 = m_0 \cdot \frac{2m_0 c^2}{W/c + m_0 c} \cdot \frac{1}{1 + \dfrac{p_z^2}{(W/c + m_0 c)^2}} \,,$$ (35)

or furthermore:

$$\bar{m}_0 = m_0 \cdot \frac{m_0 c^2}{W} = m_0 \sqrt{1 - \beta^2} \,,$$ (36)

in which βc is the velocity that corresponds to the energy W according to relativistic dynamics.

The integral invariant (33) is therefore nothing but the *action integral* (up to a constant):

$$\int^t m_0 c \sqrt{1 - \beta^2} \, dt$$ (37)

of the material point in question in relativistic dynamics. We then have the following theorem:

For a Dirac electron that is associated with a plane monochromatic wave, the time integral of the mean value of proper mass will coincide with the action integral of Einstein's dynamics.

——————

[1]   See the note on page 129.

Table of quantities and densities that are associated with the electron.

Physical quantities	Operators	Mean-value densities	Relativistic variance of the density
Proper mass (?)	$(-m_0)\,\alpha_4$	$-m_0 \sum\limits_{k=1}^{4} \Psi_k^{*}\alpha_4\Psi_k = -m_0\,\Omega_1$	Invariant
Electric charge	$-e\cdot 1$	$\delta = -\rho c = -e\sum\limits_{k=1}^{4}\Psi_k^{*}\cdot 1\cdot\Psi_k$	
Electric current	$e\,c\,\alpha_1$	$j_x = -\rho e\,u_x = e\,c\sum\limits_{k=1}^{4}\Psi_k^{*}\alpha_1\Psi_k$	Space-time vector (world current)
	$e\,c\,\alpha_2$	$j_y = -\rho e\,u_y = e\,c\sum\limits_{k=1}^{4}\Psi_k^{*}\alpha_2\Psi_k$	
	$e\,c\,\alpha_3$	$j_z = -\rho e\,u_z = e\,c\sum\limits_{k=1}^{4}\Psi_k^{*}\alpha_3\Psi_k$	
Magnetic moment $\quad M_x$	$\dfrac{eh}{4\pi m_0 c}\cdot i\,\alpha_2\alpha_3\,\alpha_4$	$I_x = \dfrac{eh}{4\pi m_0 c}\,i\sum\limits_{k=1}^{4}\Psi_k^{*}\,\alpha_2\,\alpha_3\,\alpha_4\Psi_k$	Second-rank anti-symmetric tensor Part 1
M_y	" $\cdot i\,\alpha_3\,\alpha_1\,\alpha_4$	$I_y = \dfrac{eh}{4\pi m_0 c}\,i\sum\limits_{k=1}^{4}\Psi_k^{*}\,\alpha_3\,\alpha_1\,\alpha_4\Psi_k$	
M_z	" $\cdot i\,\alpha_1\alpha_2\alpha_4$	$I_z = \dfrac{eh}{4\pi m_0 c}\,i\sum\limits_{k=1}^{4}\Psi_k^{*}\,\alpha_1\,\alpha_2\,\alpha_4\Psi_k$	

Table of quantities and densities that are associated with the electron (cont.)

Physical quantities	Operators	Mean-value densities	Relativistic variance of the density
Electric moment $\begin{cases} P_x \\ P_y \\ P_z \end{cases}$	$\dfrac{eh}{4\pi m_0 c} \cdot i\,\alpha_2\,\alpha_3$ " $\cdot\, i\,\alpha_2\,\alpha_4$ " $\cdot\, i\,\alpha_3\,\alpha_4$	$J_x = \dfrac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_1\,\alpha_4\Psi_k$ $J_y = \dfrac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_2\,\alpha_4\Psi_k$ $J_z = \dfrac{eh}{4\pi m_0 c} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_3\,\alpha_4\Psi_k$	Second-rank anti-symmetric tensor (Part 2)
Proper angular momentum $\begin{cases} S_x \\ S_y \\ S_z \end{cases}$	$\dfrac{h}{4\pi} \cdot i\,\alpha_2\,\alpha_3$ " $\cdot\, i\,\alpha_3\,\alpha_1$ " $\cdot\, i\,\alpha_1\,\alpha_2$	$\sigma_x = \dfrac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_2\,\alpha_3\Psi_k$ $\sigma_y = \dfrac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_3\,\alpha_1\Psi_k$ $\sigma_z = \dfrac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_1\,\alpha_2\Psi_k$	Completely anti-symmetric third-rank tensor = Space-time vector
(?)	S_4 " $\cdot\, i\,\alpha_1\alpha_2\,\alpha_3$	$\sigma_4 = \dfrac{h}{4\pi} i \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_1\,\alpha_2\,\alpha_3\Psi_k$	
(?)	(?) $\alpha_1\alpha_2\,\alpha_3\alpha_4$	$\Omega_2 = \sum_{k=1}^{4} \Psi_k^{*}\,\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4\Psi_k$	Completely anti-symmetric fourth-rank tensor = invariant

Table of the 16 fundamental Hermitian operators in Dirac's theory

$$\alpha_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

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

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

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

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

$$\alpha_1\alpha_2\alpha_3\alpha_4 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}$$

PART III

—

APPLICATIONS OF DIRAC'S THEORY

CRITICISMS AND VARIOUS SPECIAL TOPICS

CHAPTER XVII

EXPLANATION FOR FINE STRUCTURE
IN DIRAC'S THEORY

————

1. Wave equations for the motion of an electron in a central field ([1])

In this chapter, we propose to show that Dirac's theory provides a good explanation for the fine structure of optical and X-ray spectra that does not raise the same issues as Sommerfeld's older theory of fine structure.

Consider an electron that moves in a static central field that is derived from a potential $V(r)$. The Dirac equations for that electron will be:

$$a) \qquad \frac{2\pi i}{h}\left[\frac{W+eV}{c}+m_0 c\right]\Psi_1 - \left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\Psi_4 + \frac{\partial}{\partial z}\Psi_3 = 0,$$

$$b) \qquad \frac{2\pi i}{h}\left[\frac{W+eV}{c}+m_0 c\right]\Psi_2 - \left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\Psi_3 + \frac{\partial}{\partial z}\Psi_1 = 0,$$

$$(1)$$

$$c) \qquad \frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0 c\right]\Psi_3 - \left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\Psi_2 - \frac{\partial}{\partial z}\Psi_1 = 0,$$

$$d) \qquad \frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0 c\right]\Psi_4 - \left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\Psi_1 + \frac{\partial}{\partial z}\Psi_2 = 0.$$

It is natural to seek to express each of the Ψ_k as a product of a Laplace spherical function with a function of the radius vector.

Recall that Laplace's spherical functions have the following form:

$$Y_l^{\,m}(\theta,\varphi) = C\,e^{\pm im\varphi}\,P_l^{\,m}(\theta,\varphi) = C\,e^{\pm im\varphi}\sin^m\theta\,\frac{d^{l+m}}{(d\cos\theta)^{l+m}}(1-\cos^2\theta)^l. \quad (2)$$

Since the constant C is arbitrary, with Darwin, we choose it in such a way that we will have:

$$Y_l^{\,m}(\theta,\varphi) = (l-m)!\,e^{im\varphi}\frac{d^{l+m}}{(d\cos\theta)^{l+m}}\left(\frac{(1-\cos^2\theta)^l}{2^l\cdot l!}\right)\cdot\sin^m\theta, \qquad (3)$$

([1]) Here, we follow the method that was employed by C. G. Darwin [Proc. Roy. Soc. A **118** (1928), pp. 554].

with $l = 0, 1, 2, \ldots$, and $m = -l, -(l-1), \ldots, +l$. Moreover, we remark that the function (3) is not normalized on the surface of the sphere of unit radius.

Upon appealing to the formulas for the transformation of rectangular coordinates to polar coordinates, one can prove the following relations, in which $f(r)$ denotes an arbitrary function of the radius vector:

$$\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)f\,Y_l^{\,m}=\frac{1}{2l+1}\left\{\left[\frac{df}{dr}-\frac{l}{r}f\right]Y_{l+1}^{m+1}-(l-m)(l-m-1)\left(\frac{df}{dr}+\frac{l+1}{r}f\right)Y_{l-1}^{m+1}\right\},$$

$$\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)f\,Y_l^{\,m}=\frac{1}{2l+1}\left\{-\left[\frac{df}{dr}-\frac{l}{r}f\right]Y_{l+1}^{m-1}+(l+m)(l+m-1)\left(\frac{df}{dr}+\frac{l+1}{r}f\right)Y_{l-1}^{m-1}\right\},\quad(4)$$

$$\frac{\partial}{\partial z}f\,Y_l^{\,m}=\frac{1}{2l+1}\left\{\left(\frac{df}{dr}-\frac{l}{r}f\right)Y_{l+1}^{m}+(l+m)(l-m)\left(\frac{df}{dr}+\frac{l+1}{r}f\right)Y_{l-1}^{m}\right\}.$$

With the aid of formulas (4), we will be able to specify the derivatives that figure in (1), since we suppose that each Ψ_k is the product of a function of r with a Laplace function. We then make the hypothesis that the function Ψ_k is proportional to the function $Y_l^{\,m}$ with given values for l and m.

Upon now regarding equation (c) of (1), we will see that the terms in $Y_l^{\,m}$ that are provided by $-\dfrac{\partial\Psi_1}{\partial z}$ and $-\left(\dfrac{\partial}{\partial x}+i\dfrac{\partial}{\partial y}\right)\Psi_2$ will cancel the term in Ψ_3, and the other terms must cancel each other. It will then result that Ψ_1 and Ψ_2 must depend upon the same function of the radius vector, and must be proportional to Y_{l+1}^{m} and Y_{l+1}^{m-1}, respectively; namely, to Y_{l-1}^{m} and Y_{l-1}^{m-1}, resp.

Likewise, equation (a) of (1) shows that Ψ_1 and Ψ_2 must depend upon the same function of the radius vector and that Ψ_1 must be proportional to $Y_l^{\,m-1}$.

Finally, we are led to first imagine a solution that we can write [1]:

$$\left.\begin{array}{ll}\Psi_1=ia_1F_+(r)Y_{l+1}^{m}(\theta,\varphi), & \Psi_2=ia_2F_+(r)Y_{l+1}^{m-1}(\theta,\varphi),\\[2mm]\Psi_3=a_3G_+(r)Y_l^{\,m}(\theta,\varphi), & \Psi_4=a_4G_+(r)Y_l^{\,m-1}(\theta,\varphi).\end{array}\right\}\qquad(\mathrm{I})$$

We introduce the forms (I) into equations (a) and (b) of (1) and obtain:

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]a_1\,F_+\,Y_{l+1}^{m}$$

$$-\frac{a_1}{2l+1}\left\{\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^{m}-(l-m)(l-m+1)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m}\right\}$$

---

[1] We introduce the factor i into Ψ_1 and Ψ_2 so that a_1 and a_2 will be real. Cf., formulas (9).

$$-\frac{a_3}{2l+1}\left\{\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^m+(l+m)(l-m)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^m\right\}=0.$$

$$(5)$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]a_2\,F_+\,Y_{l+1}^{m-1}$$

$$-\frac{a_3}{2l+1}\left\{-\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^{m-1}+(l+m)(l+m-1)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m-1}\right\}$$

$$+\frac{a_4}{2l+1}\left\{\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^{m-1}+(l+m-1)(l-m+1)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m-1}\right\}=0.$$

In order to make the terms in Y_{l-1}^m disappear in the first of equations (5) and the terms in Y_{l-1}^{m-1}, in the second, it will suffice to set:

$$\frac{a_3}{a_4}=\frac{l-m+1}{l+m}.$$

$$(6)$$

On the other hand, substitute the forms (I) into equations (*c*) and (*d*) of (1); one will get:

$$\frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0c\right]a_3\,G_+\,Y_l^m$$

$$-\frac{ia_3}{2l+3}\left\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}F_+\right)Y_{l+2}^m+(?)(l-m+2)(l-m+1)\left(\frac{dF_+}{dr}+\frac{l+2}{r}F_+\right)Y_l^{\ m}\right\}$$

$$-\frac{a_1}{2l+3}\left\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}F_+\right)Y_{l+2}^m+(l+m+1)(l-m+1)\left(\frac{dF_+}{dr}+\frac{l+2}{r}F_+\right)Y_l^{\ m}\right\}=0.$$

$$(7)$$

$$\frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0c\right]a_4\,G_+\,Y_l^{m-1}$$

$$-\frac{ia_3}{2l+3}\left\{-\left(\frac{dF_+}{dr}-\frac{l+1}{r}F_+\right)Y_{l+2}^{m-1}+(l+m+1)(l+m)\left(\frac{dF_+}{dr}+\frac{l+2}{r}F_+\right)Y_l^{\ m-1}\right\}$$

$$-\frac{ia_2}{2l+3}\left\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}F_+\right)Y_{l+2}^{m-1}+(l+m)(l-m+2)\left(\frac{dF_+}{dr}+\frac{l+2}{r}F_+\right)Y_l^{\ m-1}\right\}=0.$$

In order to make the terms in Y_{l+2}^m disappear from the first equation in (7) and the terms in Y_{l+2}^{m-1}, in the second one, one must set:

$$a_4=-a_1.$$

$$(8)$$

The ratio a_1 / a_2 is arbitrary, because we can incorporate it into F_+ / G_+. We can then satisfy the conditions (6) and (8) by setting:

$$a_1 = 1, \qquad a_2 = -1, \qquad a_3 = l - m + 1, \qquad a_4 = l + m, \tag{9}$$

in such a way that $a_3 + a_4 = 2l + 1$.

Under those conditions, the two equations (5) will reduce to the single equation:

$$\frac{2\pi}{h}\left[\frac{W + eV}{c} + m_0 c\right]F_+ + \frac{dG_+}{dr} - \frac{l}{r}G_+ = 0. \tag{10}$$

Equations (7) will likewise reduce to the single equation:

$$-\frac{2\pi}{h}\left[\frac{W + eV}{c} + m_0 c\right]G_+ + \frac{dF_+}{dr} + \frac{l+2}{r}F_+ = 0. \tag{11}$$

The solution (I) is characterized by the fact that the lower indices of Y are l and $l + 1$. The mean of those two values is $l + 1 / 2$, which we call j. The solution (I) is then characterized by the quantum numbers l, $j = l + 1 / 2$, and m.

We just found a first solution in which Ψ_3 is proportional to Y_l^m. However, we have seen that there can exist another one in which Ψ_1 is proportional to Y_{l-1}^m, and Ψ_2 to Y_{l-1}^{m-1}.

We are then led to imagine a solution (II) of the form:

$$\Psi_1 = ia_1' F_-(r)Y_{l-1}^m(\theta,\varphi), \qquad\qquad \Psi_2 = ia_2' F_-(r)Y_{l-1}^{m-1}(\theta,\varphi),$$

$$\tag{II}$$

$$\Psi_2 = a_3' G_-(r)Y_l^m(\theta,\varphi), \qquad\qquad \Psi_3 = a_4' G_-(r)Y_l^{m-1}(\theta,\varphi).$$

Upon substituting the forms (II) into equations (I) and arguing as above, one will be then led to the conditions:

$$\frac{a_2'}{a_1'} = -\frac{l + m - 1}{l - m}, \qquad a_3' = -a_4', \tag{12}$$

which are analogous to (6) and (8), resp. One will satisfy those conditions by setting:

$$a_1' = -(l - m), \qquad a_2' = l + m - 1, \qquad a_3' = 1, \qquad a_4' = -1. \tag{13}$$

One will then find the following equations for F_- and G_-:

$$\frac{2\pi}{h}\left[\frac{W + eV}{c} + m_0 c\right]F_- + \frac{dG_-}{dr} + \frac{l+1}{r}G_- = 0,$$

$$\tag{14}$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_- + \frac{dF_-}{dr} - \frac{l-1}{r}F_- = 0,$$

which are the analogues of (10) and (11), resp.

The wave functions (II) constitute the solution that corresponds to the quantum numbers $l, j = l - \frac{1}{2}$, and m.

We now remark that the solution (II) will not exist for $l = 0$ (i.e., the s terms), because there are no functions Y_{-l}^m with negative indices.

We shall soon see that the determination of the functions F and G involves a quantum number n (viz., total quantum number) in such a way that any complete solution will be characterized by four quantum numbers: $n, l, j = l \pm \frac{1}{2}, m$.

We then summarize the results that were obtained in this paragraph by the following table:

Solution (I): $n, l, j = l + \frac{1}{2}, m$

$$\Psi_1 = i\, F_+\, Y_{l+1}^m, \qquad\qquad \Psi_2 = -\, i\, F_+\, Y_{l+1}^{m-1},$$
$$\Psi_3 = (l - m + 1)\, G_+\, Y_l^m, \qquad \Psi_4 = (l + m)\, G_+\, Y_l^{m-1},$$

$$\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]F_+ + \frac{dG_+}{dr} - \frac{l}{r}G_+ = 0,$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_+ + \frac{dF_+}{dr} + \frac{l+2}{r}F_+ = 0,$$

Solution (II): $n, l, j = l - \frac{1}{2}, m$

$$\Psi_1 = -\, i\, (l - m)\, F_-\, Y_{l-1}^m, \qquad \Psi_2 = i\, (l + m - 1)\, F_-\, Y_{l-1}^{m-1},$$
$$\Psi_3 = G_-\, Y_l^m, \qquad\qquad\qquad \Psi_4 = -\, G_-\, Y_l^{m-1},$$

$$\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]F_- + \frac{dG_-}{dr} + \frac{l+1}{r}G_- = 0,$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_- + \frac{dF_-}{dr} - \frac{l-1}{r}F_- = 0.$$

One remarks that the equations in this table that are satisfied by F_- and G_- can be deduced from the ones that are satisfied by F_+ and G_+ by changing l into $-(l+1)$.

2. Number of solutions that correspond to a given value of *n* **and** *l*. **Angular momentum of the electron in stationary states.** – It is very interesting to demand to know how many stationary states can exist that correspond to a given value of the quantum numbers *n* and *l*. It is obvious that there are, *a priori*, only as many states as the values of *m* that are permissible when *n* and *l* are fixed.

First, consider the solution of the type (I), and count the possible values of *m* by remarking that the upper index in the *Y* functions must be at most equal to the lower index in absolute value. There is no reason to prevent one from taking $m = 0, 1, …, l$. It might seem impossible to take $m = l + 1$, because Ψ_3 would then contain the non-existent function Y_l^{l+1}, but the coefficient a_3 is found to be zero for the value $m = l + 1$, and as a result, one can once more take that value for *m*. By contrast, the values of *m* that are greater than $l + 1$ are unacceptable.

On the other hand, we can take the negative values $- 1, - 2, …, - (l - 1)$ with no difficulty. Can it take the value $- l$? Yes, because a_4 will be zero then. However, the values of *m* that are less than $- l$ are unacceptable. Finally, for the solutions of the type (I) such that $j = l + 1 / 2$, one will have $2l + 2$ possible values for *m*, namely, *l* negative values, the value zero, and $l + 1$ positive values. In that case, there will be $2l + 2 = 2j + 1$ distinct stationary states, *a priori*.

Take the solutions of the type (II), so *m* can take the values $0, 1, …, l - 1$, and it can also take the value *l*, which makes the non-existent function Y_{l-1}^{-l} appear in Ψ_1, but it will simultaneously annul a_1'. The values of *m* that are greater than *l* or less than $- (l - 1)$ are unacceptable. Therefore, we have $2l$ possible values for *m*, namely, *l* positive values, the value zero, and $(l - 1)$ negative values.

In summary, there are $(2j + 1)$ distinct solution for the solutions (II), as there are for the solutions (I). Later, we shall see how that result permits one to justify Stoner's rule that relates to the distribution of electrons in the atom.

We shall now seek to characterize the types of solutions (I) and (II) by their corresponding angular momenta. In order to do that, we start with the result that we proved before: The first integral operator that corresponds to the angular momentum around *oz* in Dirac's theory is not $M_z = \dfrac{h}{2\pi i}\left(y\dfrac{\partial}{\partial x} - x\dfrac{\partial}{\partial y} \right) = -\dfrac{h}{2\pi i}\dfrac{\partial}{\partial \varphi}$, but in fact, $N_z =$

$M_z + S_z = -\dfrac{h}{2\pi i}\dfrac{\partial}{\partial \varphi} - \dfrac{h}{4\pi i}\alpha_1\,\alpha_2$. Now, if one remarks that $\dfrac{\partial Y_l^m}{\partial \varphi} = im Y_l^m$ then one will easily see that one will have the formulas:

$$N_z\,(\Psi_k) = -\frac{h}{2\pi i}\frac{\partial \Psi_k}{\partial \varphi} - \frac{h}{4\pi i}\alpha_1\,\alpha_2\,\Psi_k, \qquad (15)$$

for $k = 1, 2, 3, 4$, for a solution of the type (I), as well as one of type (II). One will then see that the total angular momentum N_z will be equal to $-\left(m - \tfrac{1}{2}\right)\dfrac{h}{2\pi}$ for a solution that corresponds to the quantum *m*, no matter whether it has type (I) or type (II). Now, see what the extreme values will be that this momentum can take for a given value of *l*. For a

solution of type (I), the number *m* can vary from – (*l* – 1) to + *l* for a given *l*. The momentum N_z can then vary from:

$$-\left(l-\frac{1}{2}\right)\frac{h}{2\pi} \qquad \text{to} \qquad +\left(l-\frac{1}{2}\right)\frac{h}{2\pi},$$

and a solution of type (II) will then be found to correspond to a maximum absolute value of N_z that is equal to $\left(l-\frac{1}{2}\right)\frac{h}{2\pi}$.

Now, to employ the old language, one must say that $l\dfrac{h}{2\pi}$ is the "orbital" angular momentum of the electron. We then perceive that the solutions of type (I) can be characterized by saying that they correspond to the case in which spin and orbital momentum are parallel in the same sense. On the contrary, the solutions of type (II) correspond to the case in which spin and orbital momentum are parallel in the opposite sense.

3. Calculation of the energy levels for the hydrogen atom. – We shall calculate the energy levels for the hydrogen atom. In that case, one must set $V = e / r$ in the equations of the first paragraph. We shall examine, in turn, the case of solutions of type (I), and then that of solutions of type (II).

a) Solutions of type (I):

For the solutions of type (I), we have the two equations:

$$\frac{2\pi}{h}\left[\frac{W}{c}+\frac{e^2}{cr}+m_0c\right]F_+ + \frac{dG_+}{dr}-\frac{l}{r}G_+ = 0,$$

$$-\frac{2\pi}{h}\left[\frac{W}{c}+\frac{e^2}{cr}+m_0c\right]G_+ + \frac{dF_+}{dr}-\frac{l+2}{r}F_+ = 0. \tag{16}$$

Set:

$$A^2 = \frac{2\pi}{h}\left(\frac{W}{c}+m_0c\right), \qquad B^2 = \frac{2\pi}{h}\left(\frac{W}{c}-m_0c\right), \tag{17}$$

and introduce the fine-structure constant $\alpha = \dfrac{2\pi e^2}{ch}$. We would like to find the levels of discontinuous spectra. As we know, those levels correspond to an energy of $E = W - m_0 c^2$, which is negative. The quantity B^2 will then be positive, and B will be real. With those notations, equations (16) will become:

$$\left(A^2 + \frac{\alpha}{r} \right) F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0,$$

$$\left(B^2 - \frac{\alpha}{r} \right) G_+ + \frac{dF_+}{dr} + \frac{l+2}{r} F_+ = 0. \tag{18}$$

Asymptotically, for very large r, one will have:

$$A^2 F_+ + \frac{dG_+}{dr} = 0, \qquad B^2 G_+ + \frac{dF_+}{dr} = 0, \tag{19}$$

so one will infer that:

$$\frac{d^2 G_+}{dr^2} = A^2 B^2 G_+, \qquad \frac{d^2 F_+}{dr^2} = A^2 B^2 F_+, \tag{20}$$

and upon integrating:

$$G_+ = e^{\pm ABr}, \qquad F_+ = e^{\pm ABr}. \tag{21}$$

If F_+ and G_+ are to be zero at infinity then we must take the – sign, so:

$$G_{+(\text{asyp.})} = \text{const. } e^{-ABr}, \qquad F_{+(\text{asyp.})} = \text{const. } e^{-ABr}. \tag{22}$$

These asymptotic forms suggest that we try to substitute the following forms into (18):

$$F_+ = e^{-ABr} [a_0 \, r^\gamma + a_1 \, r^{\gamma+1} + \ldots + a_s \, r^{\gamma+\sigma} + \ldots],$$

$$G_+ = e^{-ABr} [b_0 \, r^\gamma + b_1 \, r^{\gamma+1} + \ldots + b_s \, r^{\gamma+\sigma} + \ldots], \tag{23}$$

in which the exponent γ must, in principle, be assumed to be positive if F_+ and G_+ are to remain finite when $r = 0$. Upon substituting (23) into (18) and annulling the coefficients of $r^{\gamma-1}$, one will obtain:

$$\alpha \, a_0 + \gamma b_0 - l \, b_0 = 0, \qquad - \alpha b_0 + \gamma b_0 + (l+2) \, a_0 = 0. \tag{24}$$

The two equations (24) are compatible only if their determinant is zero, which will give us the condition:

$$\alpha^2 + (\gamma - l)(\gamma + l + 2) = 0. \tag{25}$$

Upon solving (25) with respect to γ, one will find:

$$\gamma = -1 + \sqrt{1 - [\alpha^2 - l(l+2)]} = -1 + \sqrt{(l+1)^2 - \alpha^2}. \tag{26}$$

We keep only the + sign in front of the radical, because the – sign will give an unacceptable negative value for γ. If we set $l = 0$ in the expression (26) then we will find a very small negative value for γ. It might seem that one should reject that value, since it corresponds to functions F_+ and G_+ (and in turn, wave functions Ψ_k) that become infinite

for $r = 0$. Now, the results to which we shall arrive later will show clearly that one must keep that solution with $l = 0$.

One can explain that anomaly by remarking that (due to the smallness of α^2) the wave functions Ψ_k will become infinite to an order at $r = 0$ that is small enough to make the integrals $\int\int\int_{-\infty}^{+\infty} |\Psi_k|^2 \, d\tau$ remain convergent. The condition that the wave functions must be everywhere finite will then seem much too rigorous here: The essential condition is that they should be square-summable. Moreover, we have seen that the same difficulty exists in the relativistic theory of one wave function. (Cf., Chap. VIII, para. 2.) We then adopt the values (26) with $l = 0, 1, 2, \ldots$ as the possible values of γ.

Having said that, if we return to the substitution of (23) into (18) and annul the coefficient of $r^{\gamma+s}$ then we will find that:

$$A \, (A \, a_s - B \, b_s) + \alpha \, a_{s+1} + (\gamma + s + 1 - l) \, b_{s+1} = 0,$$

$$- B \, (A \, a_s - B \, b_s) - \alpha \, b_{s+1} + (\gamma + s + 1 - 3) \, a_{s+1} = 0. \tag{27}$$

Multiply the first equation (27) by B, the second one, by A, and add them. One will get:

$$a_{s+1} \, [B \, \alpha + A \, (\gamma + s + l + 3)] + b_{s+1} \, [B \, (\gamma + s + l + 3) - A \, \alpha] = 0, \tag{28}$$

which will permit one to set:

$$a_{s+1} = \quad c_{s+1} \, [B \, (\gamma + s - l + 1) - A \, \alpha],$$

$$b_{s+1} = - \, c_{s+1} \, [B \, \alpha + A \, (\gamma + s - l + 3)], \tag{29}$$

in which c_{s+1} is a certain constant. Upon substituting the relations (29) into equations (27), along with the ones that one deduces by changing $s + 1$ into s, one will find the recurrence relation for the c_s :

$$A \, c_s \, [-\alpha \, (A^2 - B^2) + 2AB \, (\gamma + s + 1)] = c_{s+1} \, [A\alpha^2 + A \, (\gamma + s + 1) \, (\gamma + s - l + 1)]. \tag{30}$$

In order for the functions F_+ and G_+ to be definitely zero at infinity, from (23), it will suffice that the series $a_0 \, r + \ldots$ and $b_0 \, r + \ldots$ should be bounded. In order for that to be true, it is necessary that c_{s+1} should be zero, but not c_s , for a certain value of s, namely, $s = p$. It will then be necessary that the coefficient of c_s in (30) must be zero for $s = p$, which will give us:

$$\gamma + p + 1 = \alpha \, \frac{A^2 - B^2}{2AB}, \tag{31}$$

or, from (26):

$$\sqrt{(l+1)^2 - \alpha^2} + p = \alpha \, \frac{A^2 - B^2}{2AB}. \tag{32}$$

However, if we refer to the definitions (17) then we will see that:

$$\frac{A^2 - B^2}{2AB} = \frac{W}{\sqrt{m_0^2 - W^2}} = \frac{E + m_0 c^2}{\sqrt{-2m_0 c^2 E - E^2}} = \frac{1 + E / m_0 c^2}{\sqrt{-\frac{2E}{m_0 c^2}\left(1 + \frac{E}{2m_0 c^2}\right)}}, \tag{33}$$

in which $E = W - m_0 c^2$ is negative here. Equation (32) can then be written, after squaring:

$$\left[\sqrt{(l+1)^2 - \alpha^2} + p\right]^2 = \alpha^2 \frac{\left(1 + \frac{E}{m_0 c^2}\right)^2}{-\frac{2E}{m_0 c^2}\left(1 + \frac{E}{2m_0 c^2}\right)}, \tag{34}$$

or for that matter:

$$\frac{-\frac{2E}{m_0 c^2}\left(1 + \frac{E}{2m_0 c^2}\right)}{\left(1 + \frac{E}{m_0 c^2}\right)^2} = \frac{\alpha^2}{\left[p + \sqrt{(l+1)^2 - \alpha^2}\right]^2}. \tag{35}$$

If we add 1 to both sides of (35) then we will find:

$$\frac{1}{\left(1 + \frac{E}{m_0 c^2}\right)^2} = 1 + \frac{\alpha^2}{\left[p + \sqrt{(l+1)^2 - \alpha^2}\right]^2}, \tag{36}$$

and thus, the ultimate formula:

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{\left(p + \sqrt{(l+1)^2 - \alpha^2}\right)^2}\right]^{-1/2}. \tag{37}$$

Since one has $j = l + \frac{1}{2}$ for the solutions of the type (I) that we address here, formula (37) can also be written:

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{\left(p + \sqrt{\left(j + \frac{1}{2}\right)^2 - \alpha^2}\right)^2}\right]^{-1/2}. \tag{38}$$

Formula (38) is analogous to the Sommerfeld formula ([1]), in which the integer $j + \frac{1}{2} = l + 1$ plays the role of azimuthal quantum number, and p plays the role of radial quantum number.

As one sees, the energy levels of type (I) are defined completely by the four quantum numbers l, $j (= l + \frac{1}{2})$, m, and p. In place of the integer p that can take the values 0, 1, …, one can just as well imagine the number $n = p + l + 1$, which can take the values 1, 2, … The number n is the *total quantum number*, and the set of numbers n, l, j, m fixes the solution that one envisions.

If one develops the right-hand side of (38) up to second order in α^2 then one will get the approximate formula:

$$ E = - \frac{Rh}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) \right], \tag{39} $$

which is entirely comparable to the Sommerfeld's approximate formula ([2]), but with $j + \frac{1}{2}$, in place of the number k.

Recall that the relativistic wave mechanics of one function Ψ led us to an expression that was inconsistent with the experimental facts, which was an expression that had the same form as (38), but with l in place of j ([3]).

b) *Solutions of the type* (II).

We might repeat the calculations that we just did, while starting with the equations that the functions F_- and G_- satisfy this time. However, that would be pointless, because we remarked that we could obtain those equations by changing l into $- (l + 1)$ in equations (16). It is then obvious that upon adopting developments of the form (23) for F_- and G_-, we will get the equation for determining γ that is deduced from (26) by changing l into $- (l + 1)$, as we said; i.e.:

$$ \gamma = - 1 + \sqrt{l^2 - \alpha^2} \ . \tag{40} $$

Here, one must obviously exclude the case of $l = 0$, which will give an imaginary value: Moreover, we already know that there will be no solution of type (II) for $l = 0$. If $l = 1$ then formula (40) will give a very small negative value for γ. Here again, we assume that value of l by remarking that if the corresponding wave functions are infinite of lower order at the coordinate origin then they will nevertheless by square-summable. Briefly, we assume that l has the possible values 1, 2, …

Upon pursuing the calculations, we will obviously arrive at a formula that is deduced from (37) by the substitution of $- (l + 1)$ for l, namely:

([1]) Cf., formula (38) of Chapter I.
([2]) Cf., formula (41) of Chapter I.
([3]) Cf., formulas (34) and (36) of Chapter VIII.

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{(p + \sqrt{l^2 - \alpha^2})^2} \right]^{-1/2} . \tag{41}$$

However, for the solutions of the type (II), one will have: $j = l - \frac{1}{2}$. We will then get back to formula (38).

Instead of characterizing a solution of the type (II) by the four quantum numbers l, j ($= l - \frac{1}{2}$), m, p, we can just as well characterize them by the four numbers n, l, j, m, with $n = p + l$, and we easily recover the approximate formula (39).

In summary, in any case, the energy of the level that is characterized by the quantum numbers (n, l, j, m) is given (up to second order in α^2) by the formula (39) ([1]).

We then recover Sommerfeld's fine-structure formula for the hydrogen atom, but with the essential difference that the azimuthal number k of the old quantum theory is replaced with the integer $j + \frac{1}{2}$. It will then be between the levels whose numbers j differ by one unit that one must find the Sommerfeld doublets. The levels that differ only by the azimuthal quantum number (k or l) will coincide. That is in perfect agreement with the true fine structure of the lines in the Balmer series – for example, those of H_α. One will see that by referring to para. 5 in Chapter III, and Figure 4, in particular.

Moreover, one can rigorously repeat all of the preceding theory for a hydrogenic atom (i.e., an atom with atomic number N that has been ionized $N - 1$ times). One will easily find:

$$E = - \frac{RhN^2}{n^2} \left[1 + \frac{\alpha^2 N^2}{n^2} \left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) \right], \tag{42}$$

in place of (39).

That formula indeed accounts for the fine structure of the He^+ spectrum.

4. Application of the results obtained to the Röntgen spectra. – The formulas that are provided by Dirac's theory permit us to interpret the structure of X-ray spectra, and in particular, the existence and values of the regular doublets, without meeting up with any of the problems that were encountered in the old theory of Sommerfeld. In Chapter III, we presented the principal experimental facts that the theory had to interpret; the reader can refer to them.

We have seen that in order to calculate the energy levels in complex atoms, one can take into account (if only coarsely) the interaction of the electrons by introducing a "screening number"; viz., by replacing the atomic number N with a reduced number $N - z$ in the formulas that are valid for hydrogenic numbers. Naturally, the screening number z will vary from one electron to another, and will roughly increase as the electron gets more peripheral. Guided by the old image of the Bohr atom, we can assume that the screening number for an intra-atomic electron depends upon only the quantum numbers n and l that relate to that electron. Upon appealing to formula (42), we then write down the

---

([1]) Note that there is degeneracy since the number m does not enter into the expression for the quantized energy.

idea that the quantum numbers n, l, j, m of an electron in a complex atom have an energy that is given by the following approximate formula:

$$E\,(n,\,l,\,j,\,m) = -\,\frac{Rh(N - z_{nl})^2}{n^2}\left[1 + \frac{\alpha^2(N - z_{nl})^2}{n^2}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right)\right].$$

(43)

That formula differs from the old Sommerfeld formula [cf., Chap. III, form. (9)] by the substitution of $j + \frac{1}{2}$ for k. However, that simple substitution will suffice to allow one to interpret the regular doublets with formula (43) without lending itself to the objections that were raised by the old Sommerfeld formula. Indeed, it is obvious that the regular doublets can be predicted by formula (43), not between levels with azimuthal quantum numbers (k or l) that differ by one unit, but between levels with the same azimuthal quantum number and j numbers that differ by one unit. That is, in fact, what experiments demand, as we have seen.

For example, consider the levels L_{II} and L_{III}, whose frequency difference corresponds to the Sommerfeld doublets. We know that:

$$\text{for } L_{II}: \qquad n = 2, \;\; l = 1, \;\; j = \tfrac{1}{2},$$
$$\text{for } L_{III}: \qquad n = 2, \;\; l = 1, \;\; j = \tfrac{3}{2}.$$

We then find from (43) that the frequency difference of the corresponding doublets is:

$$\nu_{L_{II}} - \nu_{L_{III}} = -\,\alpha^2\,\frac{R(N - z_{21})^4}{2^2}\left[\frac{2}{\frac{1}{2}+\frac{1}{2}} - \frac{2}{\frac{3}{2}+\frac{1}{2}}\right] = -\,\frac{R\alpha^2}{16}\,(N - z_{21})^4,$$

(44)

and we have seen that upon setting $z = 3.5$, we will get good agreement with experiments, and here, the doublet is predicted to be in its true place between two levels with the same l, and j's that differ by one unit.

We have pointed out that there also exist irregular doublets in the X-ray spectra. The origin of those doublets is the difference in frequency that exists between two neighboring levels with the same quantum number j and l numbers that differ by one unit. We can also explain their existence by using formula (43). Indeed, if the screening number z does not depend upon l then the levels with the same j and different l (such as L_I and L_{II}, for example) will coincide: That is what happens with hydrogen, because, of course, z is always zero in that case. However, by reason of the variation of z with l, one will see that those levels must not coincide in complex atoms, and one can even predict a law for their frequency difference. Indeed, if one neglects the terms of order higher than α^2 then one will have:

$$\sqrt{\frac{E(n,l,\,j)}{h}} = \sqrt{\frac{Rh}{n^2}}\,(N - z_{n.l}),$$

$$\sqrt{\frac{E(n,l+1,\,j)}{h}} = \sqrt{\frac{Rh}{n^2}}\,(N - z_{n.l+1}).$$

(45)

Hence:

$$\sqrt{\frac{E(n,l,\,j)}{h}} - \sqrt{\frac{E(n,l+1,\,j)}{h}} = \sqrt{\frac{Rh}{n^2}}\,(z_{n,\,l+1} - z_{n,\,l}).\qquad(46)$$

The left-hand side of (46) is $\delta\sqrt{v}$ for the two levels, while the right-hand side is independent of N. One will then see that the spectral terms for the two levels in the series of elements are such that $\delta\sqrt{v}$ is independent of N. That is, in fact, the law of irregular doublets that was pointed out at the end of para. 2 in Chapter III. Those irregular doublets exist between spectral terms that are characterized by the same letter, but with indices that differ by one unit, *the first of which is odd*, while Sommerfeld's regular doublets provide spectral terms that are characterized by the same letter and indices that differ by one unit, *the first of which is even*. Hence, the difference $v_{L_I} - v_{L_{II}}$ will give rise to irregular doublets, and the difference $v_{L_{II}} - v_{L_{III}}$ will give rise to regular ones.

Dirac's theory has then given us, and in a good form, the entire Sommerfeld theory of fine structure and has shown us that its original success was not due to chance, as well as showing us that, in the final analysis, the existence of regular doublets is indeed linked with relativity by the intermediary of electron spin.

5. Number of electrons per level. Stoner's rule. – Dirac's theory has not only allowed us to recover the fine-structure formulas, while improving upon them, it has also provided an interpretation for Stoner's rule that relates to the distribution of electrons between the atomic levels and a proof of the selection rules that relate to the quantum numbers l, m, and j.

Here, we shall first address the first question and recall that Stoner's rule is the following one: *One cannot have more than* $2j + 1$ *electrons in the energy level that corresponds to the quantum numbers n, l, j.* We shall seek to justify that statement.

We saw above (para. 2) that there exist $2j + 1$ levels that correspond to the set of quantum numbers (n, l, j) for solutions of type (I), as well as for solutions of type (II). In other words, since each level is defined completely by four quantum numbers n, l, j, m, there will exist $2j + 1$ possible values for m, and in turn, $2j + 1$ possible levels for given values of the three numbers n, l, $j = l \pm \frac{1}{2}$. Having recalled that, in order to go further, one must appeal to a new principle that plays a fundamental role in contemporary physics: the Pauli Exclusion Principle.

We shall state the Pauli Exclusion Principle here by saying: *One cannot have more than one intra-atomic electron whose stationary state is characterized by the same four integers n, l, j, m.*

Now, as we have seen, the energy of a stationary state in an atom in the absence of an external field will not depend upon the number m, in such a way that each energy level is characterized entirely by the three quantum numbers n, l, j, which is what experiments indicated before the theory, moreover (cf., Part I). Hence, if one assumes the Pauli Exclusion Principle then the number of electrons that belong to the same level (n, l, j) is equal to at most the number of stationary solutions that are characterized by the four quantum numbers n, j, l, m, the first three of which characterize the level under

consideration. We know that this number of solutions is $2j + 1$. We have then justified Stoner's rule, whose experimental exactitude does not seem to be in doubt.

6. Selection rules in Dirac's theory. – In paragraph 4 of Chapter VIII, we explained how the correspondence principle leads one to predict selection rules in wave mechanics. In Dirac's theory, that prediction is made in the same fashion, but while taking into account the fact that the expression $\Psi\Psi^*$ in wave mechanics with one wave function is replaced with:

$$\sum_{k=1}^{4} \Psi_k^* \Psi_k$$

here.

One will then have to consider the following quantities:

$$X_{nm} = \int\int\int_{-\infty}^{+\infty} x(\Psi_{1,n}^*\Psi_{1,m} + \Psi_{2,n}^*\Psi_{2,m} + \Psi_{3,n}^*\Psi_{3,m} + \Psi_{4,n}^*\Psi_{4,m})\,d\tau \qquad (47)$$

to be the X matrix elements that serve to evaluate the transition probabilities that accompany radiation. The indices n and m characterize two stationary states, and in reality, they will each represent four indices n, l, j, m. If the elements X_{nm}, Y_{nm}, Z_{nm} are zero then there will be no radiation that corresponds to the transition $n \to m$. One infers the selection rules from that.

As an example, consider the transition from the stationary state of type (I) with quantum numbers n, l, $j = l + \frac{1}{2}$, m to the stationary state of type (II) with quantum numbers n, $l - 1$, $j = l - \frac{3}{2}$, m, which is a transition for which one will have $| \delta l | = 1$, $\delta m = 0$, and $| \delta j | = 2$. The first of those states has the wave functions:

$$\Psi_1 = i\,F_+(r)\,Y_{l+1}^m, \qquad\qquad \Psi_2 = -\,i\,F_+(r)\,Y_{l-1}^{m-1},$$
$$\Psi_3 = (l - m + 1)\,G_+(r)\,Y_l^m, \qquad \Psi_4 = (l + m)\,G_+(r)\,Y_l^{m-1},$$

and the second one has:

$$\Psi_1 = -\,i\,(l - m - 1)\,F_-(r)\,Y_{l-2}^m, \qquad \Psi_2 = i\,(l + m - 2)\,F_-(r)\,Y_{l-2}^{m-1},$$
$$\Psi_3 = G_-(r)\,Y_{l-1}^m, \qquad\qquad \Psi_4 = -\,G_-(r)\,Y_{l-1}^{m-1}.$$

We then write the expression for the Z-matrix element that corresponds to that transition, by setting:

$$A = -\int_0^{+\infty} F_+(r)\,F_-(r)\,r^2 dr, \quad B = \int_0^{+\infty} G_+(r)\,G_-(r)\,r^2 dr, \qquad (48)$$

to simplify.

One will get:

$$Z_{nm} = A \, (l - m - 1) \iint \cos\theta \, Y_{l-2}^{m*} \, Y_{l+1}^{m} \, d\Omega \; + A \, (l + m - 2) \iint \cos\theta \, Y_{l-2}^{m-1*} \, Y_{l+1}^{m-1} \, d\Omega$$

$$+ \, B \, (l - m + 1) \iint \cos\theta \, Y_{l-1}^{m*} \, Y_{l}^{m} \, d\Omega \; - B \, (l + m) \iint \cos\theta \, Y_{l-1}^{m-1*} \, Y_{l}^{m-1} \, d\Omega, \qquad (49)$$

with $d\Omega = \sin\theta \, d\theta \, d\varphi$, and the double integrals are extended over the surface of the sphere of *unit* radius. Now, if one starts with the definition of the Y functions then one can prove the following formulas:

$$\iint \cos\theta \, Y_{l-2}^{m*} \, Y_{l+1}^{m} \, d\Omega \quad = 0,$$

$$\iint \cos\theta \, Y_{l-1}^{m*} \, Y_{l}^{m} \, d\Omega \quad = \frac{4\pi}{(2l+1)(2l-1)} \, (l+m)! \, (l-m)!,$$

$$\iint \cos\theta \, Y_{l-2}^{m-1*} \, Y_{l+1}^{m-1} \, d\Omega = 0,$$

$$\iint \cos\theta \, Y_{l-1}^{m-1*} \, Y_{l}^{m-1} \, d\Omega = \frac{4\pi}{(2l+1)(2l-1)} \, (l+m-1)! \, (l-m)!$$

If one starts with (49) then one will find that:

$$Z_{nm} = B \, \frac{4\pi}{(2l+1)(2l-1)}$$

$$\times \, [(l-m+1) \, (l+m)! \, (l-m)! - (l+m) \, (l-m+1)! \, (l+m-1)!] = 0. \qquad (51)$$

One will likewise find that $X_{nm} = Y_{nm} = 0$. The transition envisioned is then forbidden.

One can repeat the same type of calculations by taking all combinations of a solution of type (I) and a solution of the same type or of type (II), and then all combinations of a solution of type (II) with a solution of the same type or of type (II). The result of those calculations will be as follows:

The only transitions that correspond to radiation are the ones for which one has:

$$\delta l = \pm 1, \qquad \delta m = \left\{ \begin{array}{c} \pm 1 \\ 0 \end{array} \right., \qquad \delta j = \left\{ \begin{array}{c} \pm 1 \\ 0 \end{array} \right..$$

Those selection rules are, in fact, the ones that one found empirically. The rules that relate to the quantum numbers l and m are predicted by wave mechanics with just one wave function, but the rule that relates to the quantum number j can only be predicted by a theory that introduces that quantum number, as is the case for Dirac's theory.

_____

CHAPTER XVIII

DEDUCTION OF THE LANDÉ FORMULA
____

1. Summary of the method of perturbations

In this chapter, we propose to show that Dirac's theory permits us to recover the Landé formula for the anomalous Zeeman Effect of the alkali metals in a weak magnetic field. However, since we must make use of the theory of perturbations in order to do that, we shall first say a few words about that method of calculating.

Suppose that we have determined the stationary states of a system – a hydrogen atom, for example. We then know the proper values W_n of its energy and the corresponding proper functions $\Psi_{k,n}$, in which k always denotes the Dirac index, namely, the *spin* variable. The symbolic equation:

$$\left[\frac{W_n + eV}{c} + \sum_{j=1}^{3} \alpha_j P_j + \alpha_4 m_0 c \right] \Psi_n = 0 \tag{1}$$

will then be satisfied when the system is not perturbed.

However, now suppose that a slight, constant, perturbing action is exerted upon the quantized system and that it can be represented by the addition of a term $\Lambda\Psi$ to equation (1). (Λ is an operator that can contain the α_i and, in turn, operate on the index k.) As a result of the presence of the small perturbing term, the proper values and functions must be slightly modified, and will become $W_n + \varepsilon_n$ and $\Psi_{k,n} + \eta_{k,n}$, resp.

We suppose that the ε_n and the $\eta_{k,n}$ are very small, like the perturbing term $\Lambda\Psi$, and we neglect terms such as $\varepsilon_n \eta_{k,n}$ and $\Lambda\eta_{k,n}$.

In the perturbed state, the symbolic equation will be:

$$\left[\frac{W_n + \varepsilon_n + eV}{c} + \sum_{j=1}^{3} \alpha_j P_j + \alpha_4 m_0 c + \Lambda \right] (\Psi_n + \eta_n) = 0. \tag{2}$$

Subtract (1) from (2); with the assumed approximations, we will then find that:

$$\left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_n + \left[\frac{W_n + eV}{c} + \sum_{j=1}^{3} \alpha_j P_j + \alpha_4 m_0 c \right] \eta_n = 0. \tag{3}$$

Now consider the four functions $\eta_{k,n}$. It results from formula (5) of chapter XVI that we can develop those functions in a complete system of functions Ψ_n by means of the formulas:

$$\eta_{k,n} = \sum_m c_{n,m} \Psi_{k,m} \qquad (k = 1, 2, 3, 4), \qquad (4)$$

or symbolically:

$$\eta_n = \sum_m c_{n,m} \Psi_m . \qquad (5)$$

If one takes into account the equation that one obtains by changing n into m in (1) then one will have:

$$\left(\frac{W_n + eV}{c} + \sum_{j=1}^3 \alpha_j P_j + \alpha_4 m_0 c \right) \eta_n$$

$$= \sum_m c_{n,m} \left[\frac{W_n + eV}{c} + \sum_{j=1}^3 \alpha_j P_j + \alpha_4 m_0 c \right] \Psi_n$$

$$= \sum_m c_{n,m} \frac{W_n - W_m}{c} \Psi_m , \qquad (6)$$

and, in turn, from (3):

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \Psi_m = \left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_n , \qquad (7)$$

or, upon making that explicit:

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \Psi_{k,m} = \left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_{k,n} . \qquad (8)$$

Multiply (8) by $\Psi_{k,m}^*$, sum over k, and integrate over all space. One will get:

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \int\limits_{-\infty}^{+\infty}\!\!\int\int \sum_k \Psi_{k,l}^* \Psi_{k,m} \, d\tau = \int\limits_{-\infty}^{+\infty}\!\!\int\int \sum_k \Psi_{k,l}^* \left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_{k,n} \, d\tau . \qquad (9)$$

If one then takes into account the orthogonality and normalization of the proper functions then one will infer from this that:

$$c_{n,l} = \frac{c}{W_n - W_l} \int\limits_{-\infty}^{+\infty}\!\!\int\int \sum_k \Psi_{k,l}^* \left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_{k,n} \, d\tau . \qquad (10)$$

That formula will give the c_{nl} for all values of l.

However, formula (10) will give us an infinite coefficient c_{nn} for $l = n$ when the integral that appears in (10) is non-zero for $l = n$. Since that is not acceptable, we see that we must have:

$$\int\!\!\int_{-\infty}^{+\infty}\!\!\int \sum_k \Psi_{k,n}^* \left(\frac{\varepsilon_n}{c} + \Lambda \right) \Psi_{k,n}\, d\tau = 0. \tag{11}$$

The condition (11), which expresses a well-known theorem of Fredholm in the theory of integral equations, permits one to obtain the variation ε_n of the energy of the nth stationary state when it is provoked by the presence of the perturbation. Indeed, one infers from (11) that:

$$\frac{\varepsilon_n}{c} = - \int\!\!\int_{-\infty}^{+\infty}\!\!\int \sum_k \Psi_{k,n}^* \Psi_{k,n}\, d\tau , \tag{12}$$

because the $\Psi_{k,n}$ are normalized.

It should be remarked that formula (10) is acceptable only if it leads to very small values of c_{nl} . Otherwise, the hypothesis of the smallness of the η_n to which we have appealed will not be valid. Upon examining that point, we will see that formula (12) is exact only if we can (as we have supposed implicitly) consider the perturbation of each proper value as being independent of the ones that the other proper values are subjected to. For that to be true, the external perturbations must be small enough that the displacement of proper values that is due to perturbation will be small compared to the difference between those proper values. That is what happens with the Zeeman Effect when the magnetic field is weak. If that condition is not satisfied (the Zeeman Effect with a strong magnetic field, for example) then there will good reasons to conduct the calculation of the perturbations in a somewhat different way; we shall not elaborate upon that here.

2. Application of formula (12) in the case of the Zeeman Effect. – In Chapter IV, we saw how the Landé formula permitted us to represent the modification of an energy level in an atom with a doublet spectrum (e.g., an alkali metal) by the action of a weak magnetic field. If $W_0\,(n, l, j)$ is the energy of the level in the absence of the magnetic field, and if $W_H\,(n, l, j)$ is that energy in the presence of the field H then one will have:

$$W_H\,(n, l, j) = W_0\,(n, l, j) + m'g\,\frac{eh\,H}{4\pi m_0 c}, \tag{13}$$

with

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}} = \frac{2j+1}{2l+1}. \tag{14}$$

The number m' is a positive or negative half-integer number (i.e., an integer plus $\frac{1}{2}$) that will take on all half-integer values from $-j$ to $+j$. We have modified the notations of Chapter IV by writing m', instead of m, in order to distinguish the half-integer number m' in formula (13) from the fourth quantum number m that characterizes a stationary state of the atom, and which is an integer. We see that the numbers m' and m are united by the relation $m' = -\left(m - \frac{1}{2}\right)$. Let us make that point more precise: In formula (13), we define the energy level by its three quantum numbers n, l, j, which will suffice since the value of

the energy does not depend upon the fourth quantum number m.   However, the displacement of the level by the Zeeman Effect will depend upon m [1], and upon making that dependency explicit, we can write (13) in the form:

$$W_H\,(n,\,l,\,j,\,m) = W_0\,(n,\,l,\,j) - \left(m - \tfrac{1}{2}\right)\frac{2j+1}{2l+1}\frac{eh\,H}{4\pi m_0 c}\,. \tag{15}$$

That is the formula that Dirac's theory will permit us to recover.

We consider an atom, in which it is permissible to imagine that it has just one electron: That is rigorously the case for a hydrogen atom of rank N that has been ionized $(N-1)$ times.   It is approximately the case for an alkali metal atom in which the valence electron is unique and in which one can roughly account for the presence of other peripheral electrons by a simple screening effect.   We suppose that this atom is embedded in a uniform magnetic field **H** that is directed along the z-axis.   We can then take the potential vector to be:

$$A_x = -\tfrac{1}{2}y\,H, \quad A_y = \tfrac{1}{2}x\,H, \qquad A_z = 0. \tag{16}$$

The perturbing term in equation (2) is then:

$$\Lambda\Psi = \frac{e}{c}(\alpha_1\,A_x + \alpha_2\,A_y + \alpha_3\,A_z)\,\Psi = \frac{eH}{2c}(x\,\alpha_2 - y\,\alpha_1)\,\Psi. \tag{17}$$

Upon expanding this by using the well-known values for α_1 and α_3, one will find:

$$\Lambda\Psi_1 = i\,(x+i\,y)\,\frac{eH}{2c}\,\Psi_4\,, \qquad\qquad \Lambda\Psi_2 = -\,i\,(x-i\,y)\,\frac{eH}{2c}\,\Psi_3\,,$$

$$\tag{18}$$

$$\Lambda\Psi_3 = i\,(x+i\,y)\,\frac{eH}{2c}\,\Psi_2\,, \qquad\qquad \Lambda\Psi_4 = -\,i\,(x-i\,y)\,\frac{eH}{2c}\,\Psi_1\,.$$

An application of formula (12) then gives us the displacement of the level by the Zeeman Effect in a weak magnetic field as:

$$\frac{\varepsilon}{c} = \frac{W_H\,(n,l,\,j,m) - W_0(n,l,\,j)}{c}$$

$$= \frac{eH}{2c}\int\int_{-\infty}^{+\infty}\!\!\!\int [-\Psi_1^{*}i(x+iy)\Psi_4 + \Psi_2^{*}i(x-iy)\Psi_3 - \Psi_3^{*}i(x+iy)\Psi_2 + \Psi_4^{*}i(x-iy)\Psi_1]\,d\tau. \tag{19}$$

The entire problem is now to calculate the integral that appears in (19).

---

[1]   One can say that the presence of the magnetic field made the degeneracy disappear that existed in its absence.

3. Development of the calculations. – In order to perform the calculation, we must distinguish two cases, according to whether the stationary state of energy W_0 (n, l, j) has type (I) or type (II).

a) *Solution of type* (I):

$$\Psi_1 = i\, F_+\, (r)\, Y_{l+1}^m, \qquad\qquad \Psi_2 = -\, i\, F_+\, (r)\, Y_{l+1}^{m-1},$$

$$\Psi_3 = (l - m + 1)\, G_+\, (r)\, Y_l^m, \quad \Psi_4 = (l + m)\, G_+\, (r)\, Y_l^{m-1}, \tag{20}$$

in which F_+ and G_+ are two real functions of r that obey simultaneous equations:

$$\frac{2\pi}{h}\left[\frac{W_0 + eV}{c} + m_0 c\right] F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0,$$

$$-\frac{2\pi}{h}\left[\frac{W_0 + eV}{c} - m_0 c\right] G_+ + \frac{dF_+}{dr} + \frac{l+2}{r} F_+ = 0. \tag{21}$$

As in the atom, the Newtonian approximation is always valid approximately, so the functions Ψ_1 and Ψ_2 will be very small with respect to Ψ_3 and Ψ_4 ; the ratio F_+ / G_+ will then be very small. A more complete calculation proves that this ratio has order the fine-structure constant α. We appeal to that fact in order to neglect the square $(F_+ / G_+)^2$, which will afford some notable simplifications.

Upon multiplying the first of equations (21) by G_+ and the second one by F_+ and adding them, one will infer that:

$$\frac{4\pi\, m_0 c}{h}\, F_+ G_+ + \frac{1}{2}\frac{d}{dr}\left[F_+^2 + G_+^2\right] + \frac{l+2}{r} F_+^2 - \frac{l}{r} G_+^2 = 0. \tag{22}$$

Neglect the terms in F_+, multiply by $r^2\, dr$, and integrate from 0 to $+\infty$; one will get:

$$\int_0^{+\infty} F_+ G_+\, r^3 dr = -\,\frac{h}{4\pi\, m_0 c}\int_0^{+\infty}\left[\frac{r^2}{2}\frac{dG^2}{dr} - l\, G_+^2 r^2\right] dr. \tag{23}$$

One will get G_+, which is zero at infinity, upon integrating this by parts:

$$\int_0^{+\infty} F_+ G_+\, r^3 dr = \frac{h}{4\pi\, m_0 c}\left(l + \frac{3}{2}\right)\int_0^{+\infty} G_+^2\, r^2\, dr. \tag{24}$$

Endowed with those formulas, we can now begin the calculation of the integral that appears in the right-hand side of (19). Upon taking polar coordinates, we will have:

$$x + i\, y = r \sin\theta\, e^{i\varphi}, \quad x - i\, y = r \sin\theta\, e^{-i\varphi}, \quad d\tau = r^2 \sin\theta\, dr\, d\theta\, d\varphi. \tag{25}$$

Hence:

$$\int\int\int_{-\infty}^{+\infty} \left[-\Psi_1^* i(x+iy)\Psi_4 + \Psi_2^* i(x-iy)\Psi_3 - \Psi_3^* i(x+iy)\Psi_2 + \Psi_4^* i(x-iy)\Psi_1 \right] d\tau$$

$$= - \int_0^{2\pi}\int_0^{\pi} \left[Y_{l+1}^{m*}(l+m) Y_l^{m-1} e^{i\varphi} + Y_{l+1}^{m-1*}(l-m+1) Y_l^m e^{-i\varphi} + (l-m+1) Y_l^{m*} Y_{l+1}^{m-1} e^{i\varphi} \right.$$

$$\left. (l+m) Y_l^{m-1*} Y_{l+1}^m e^{-i\varphi} \right] \sin^2\theta\, d\theta\, d\varphi \int_0^{+\infty} F_+ G_+\, r^2 dr \,. \tag{26}$$

Now, a study of the Y functions will yield the integrals:

$$\int_0^{2\pi}\int_0^{\pi} Y_{l+1}^{m-1*} Y_l^{m-1} e^{i\varphi} \sin^2\theta\, d\varphi = \frac{4\pi}{(2l+1)(2l+3)}(l+m+1)!(l-m+1)!,$$

$$\int_0^{2\pi}\int_0^{\pi} Y_{l+1}^{m-1*} Y_l^m\, e^{-i\varphi} \sin^2\theta\, d\varphi = -\frac{4\pi}{(2l+1)(2l+3)}(l+m)!(l-m+2)!,$$

$$\tag{27}$$

$$\int_0^{2\pi}\int_0^{\pi} Y_l^{m*} Y_{l+1}^{m-1} e^{i\varphi} \sin^2\theta\, d\varphi = -\frac{4\pi}{(2l+1)(2l+3)}(l+m)!(l-m+2)!,$$

$$\int_0^{2\pi}\int_0^{\pi} Y_l^{m-1*} Y_{l+1}^m\, e^{-i\varphi} \sin^2\theta\, d\varphi = \frac{4\pi}{(2l+1)(2l+3)}(l+m+1)!(l-m+1)!.$$

One will then find the following value for the integral (26):

$$-\frac{4\pi}{(2l+1)(2l+3)}(l+m)!\,(l-m+1)!$$

$$\times 2\left[(l+m)(l+m+1)-(l-m+1)(l-m+2)\right] \int_0^{+\infty} F_+ G_+\, r^2 dr$$

$$\tag{28}$$

$$= -\frac{4\pi}{2l+1}(l+m)!\,(l-m+1)!$$

$$\times \left[(l+m)(l-m+1)-(l-m+1)(l-m+2)\right] \frac{h}{4\pi m_0 c} \int_0^{+\infty} G_+^2\, r^2 dr \,,$$

in which the last term is obtained by introducing the expression (24).

Since one can easily verify the following identity:

$$(l+m)(l+m+1) - (l\text{-}m+1)(l-m+2) = (2l+2)(2m-1), \tag{29}$$

one can further write the expression (28) in the form:

$$-4\pi \cdot \frac{2l+2}{2l+1}(2m-1)(l+m)!\,(l-m+1)!\,\frac{h}{4\pi\,m_0 c}\int_0^{+\infty} G_+^2\,r^2 dr\,. \tag{30}$$

In order to go further, one must invoke the normalization condition, and upon neglecting the terms in F_+^2 in comparison to the ones in G_+^2, that will give:

$$\int_0^{2\pi}\int_0^{\pi}\Big[(l-m+1)^2 Y_l^{m*}\,Y_l^m + (l+m)^2 Y_l^{m-1*}\,Y_l^{m-1}\Big]\sin\theta\,d\theta\,d\varphi\int_0^{+\infty} G_+^2 r^2 dr = 1. \tag{31}$$

Now, one has:

$$\int_0^{2\pi}\int_0^{\pi}\big|Y_l^m\big|^2\sin\theta\,d\theta\,d\varphi = \frac{4\pi}{2l+1}(l+m)!\,(l-m)!,$$

$$\int_0^{2\pi}\int_0^{\pi}\big|Y_l^{m-1}\big|^2\sin\theta\,d\theta\,d\varphi = \frac{4\pi}{2l+1}(l+m-1)!\,(l-m+1)!, \tag{32}$$

in such a way that the condition (31) will give us:

$$\int_0^{+\infty} G_+^2\,r^2 dr = \frac{1}{4\pi(l+m)!(l-m+1)!}\,. \tag{33}$$

Upon substituting that value into the expression (30), we will finally see that the integral to be calculated in formula (19) has simply the value:

$$-\frac{2l+2}{2l+1}(2m-1)\frac{h}{4\pi\,m_0 c}\,. \tag{34}$$

Formula (19) will then provide us with the following formula for the level displacement by the Zeeman Effect:

$$W_H\,(n,\,l,\,j,\,m) = W_0\,(n,\,l,\,j) - \frac{eH}{2}\frac{2l+2}{2l+1}(2m-1)\frac{h}{4\pi\,m_0 c}\,. \tag{35}$$

Now set:

$$m' = -\left(m-\tfrac{1}{2}\right), \qquad g = \frac{l+1}{l+\tfrac{1}{2}} = \frac{j+\tfrac{1}{2}}{l+\tfrac{1}{2}}, \tag{36}$$

while remembering that $j = l + \tfrac{1}{2}$ here. We can write:

$$W_H\,(n,\,l,\,j,\,m) = W_0\,(n,\,l,\,j) + m'\,g\,\frac{ehH}{4\pi\,m_0 c}\,. \tag{37}$$

If we now suppose only that we are given the three quantum numbers n, l, j then we will have:

$$W_H (n, l, j) = W_0 (n, l, j) + m' \, g \, \frac{ehH}{4\pi \, m_0 c}. \tag{38}$$

Here, m' is a half-integer that can vary from $-\left(m_{\min} - \frac{1}{2}\right)$ to $-\left(m_{\max} - \frac{1}{2}\right)$. Now, when we discussed the number of levels that correspond to the triple of quantum numbers (n, j, l), we saw that for a solution of type (I), one can make m vary from $-l$ to $l + 1$. For a given (n, j, l), m' can vary from $-\left(-l - \frac{1}{2}\right) = l + \frac{1}{2} = j$ to $-\left(l + 1 - \frac{1}{2}\right) = -\left(l + \frac{1}{2}\right) = -j$.

Briefly, we have formula (38) for the levels that correspond to solutions of type (I), in which m' is a half-integer that can vary from $-j$ to $+j$. That is indeed the Landé formula.

b) *Solution of type* (II):

In order to carry out the calculation in the case of a solution of type (II), one starts with the wave function:

$$\Psi_1 = - i \, (l - m) \, F_- \, Y_{l-1}^m, \qquad \Psi_2 = i \, (l + m - 1) \, F_- \, Y_{l-1}^{m-1},$$

$$\tag{39}$$

$$\Psi_3 = G_- \, Y_l^m, \qquad\qquad \Psi_4 = - G_- \, Y_l^{m-1},$$

in which the functions F_- and G_- are real functions that satisfy the equations:

$$\frac{2\pi}{h} \left[\frac{W + eV}{c} + m_0 c \right] F_- + \frac{dG_-}{dr} + \frac{l+1}{r} G_- = 0,$$

$$\tag{40}$$

$$- \frac{2\pi}{h} \left[\frac{W + eV}{c} - m_0 c \right] G_- + \frac{dF_-}{dr} - \frac{l-1}{2} F_- = 0.$$

One will arrive at the following formula by calculations that are analogous to the ones that were developed above:

$$W_H (n, l, j, m) = W_0 (n, l, j) - \frac{2l}{2l+1} (2m - 1) \frac{eH}{2} \cdot \frac{h}{4\pi \, m_0 c}. \tag{41}$$

Since we have $j = l - \frac{1}{2}$ for the solutions (II), we set:

$$m' = -\left(m - \frac{1}{2}\right), \qquad g = \frac{l}{l + \frac{1}{2}} = \frac{j + \frac{1}{2}}{l + \frac{1}{2}}, \tag{42}$$

and we can write (41) in the form:

$$W_H\,(n,\,l,\,j) = W_0\,(n,\,l,\,j) + m'\;g\;\frac{eh\,H}{4\pi\,m_0 c}\,,\qquad(44)$$

in which m' is a half-integer that varies from $-\left(m_{\min} - \tfrac{1}{2}\right)$ to $-\left(m_{\max} - \tfrac{1}{2}\right)$. Now (as we have seen), we can make m vary from $-(l-1)$ up to $+l$, here. The number m can then vary from $-\left[-(l-1)-\tfrac{1}{2}\right]= l-\tfrac{1}{2}$ to $-\left(l-\tfrac{1}{2}\right)$; i.e., from $+j$ to $-j$.

Briefly, we have formula (44) for the levels that correspond to type (II), in which m' is a half-integer that can vary from $-j$ to $+j$.

4. Summary of results. Remarks. – We have thus proved that for all of the levels [of type (II), as well as type (I)], the displacement of the level $(n,\,l,\,j)$ by the action of a sufficiently-weak, uniform magnetic field is given by:

$$W_H\,(n,\,l,\,j) - W_0\,(n,\,l,\,j) = m'\;g\;\frac{eh\,H}{4\pi\,m_0 c}\,,\qquad(45)$$

with:

$$g = \frac{j+\tfrac{1}{2}}{l+\tfrac{1}{2}}\,,\qquad(46)$$

in which m' is a half-integer that can take all half-integer values between $-j$ and $+j$. That is, in fact, the anomalous Zeeman Effect, as it is represented exactly by the Landé formula. Since the number m' has $2j + 1$ possible values, any level $(n,\,l,\,j)$ will be decomposed by the Zeeman Effect into $(2j + 1)$ distinct elements, which can be expressed by saying the external magnetic field has made the degeneracy of order $2j + 1$ that existed in its absence disappear.

The argument that we just made is rigorously valid for hydrogenic atoms, but as we have pointed out, one can also apply it to alkali metal atoms, at least approximately.

We have seen that the Landé formula is valid for weak magnetic fields. By that, we mean magnetic fields that are weak enough that the displacement of the energy levels by the Zeeman Effect is small with respect to the normal distance between those levels. If that condition is not satisfied then formula (12) will no longer be valid, and we must repeat all of the calculations. It can be proved [1] that we will then arrive at a formula that is identical to that of Voigt [cf., Chap. IV, end of para. 3], and for very strong fields, we will find the Paschen-Back effect; i.e., the normal Zeeman decomposition.

Dirac's theory then gives us a completely satisfactory solution to the anomalous Zeeman Effect for the alkali metals. Dirac's theory cannot rigorously predict the Zeeman Effect for atoms for which it is not possible to consider just one optical electron, because that theory does not further know how to treat the case of a system of interacting electrons.

—————

[1] See C. G. DARWIN, Proc. Roy. Soc. **118** (1928), pp. 654.

PROPER AND ORBITAL MOMENTS. POLARIZATION OF ELECTRONIC WAVES

1. Impossibility of separating the proper magnetic moment from the orbital magnetic moment (Bohr).

Bohr showed by a very delicate argument that it cannot be possible to measure the proper magnetic moment of an electron, because the effects that are due to the existence of that proper moment cannot be distinguished from the ones that are due to the collective motion of the electron.

We shall summarize Bohr's argument.

In order to exhibit the proper magnetic moment of the electron, one can proceed in two ways:

1. Seek to measure the effect of a small magnet that is equivalent to the electron on a magnetometer.

2. Make the electron traverse a non-uniform magnetic field, and observe the effect of that field on the small magnet.

Let us examine the first method. If the direction of the motion of the electron is taken to be the x-axis then we will place the magnetometer on the y-axis at the point whose ordinate is y. In order to be able to write the expression for its effect on the magnetometer with any precision, we must suppose that the electron is sufficiently localized; i.e., that the electron must be associated with a wave train Ψ whose dimensions are small with respect to the distance y from the magnetometer to the axis of the motion. For example, if Δx is the length of that wave train along the x-axis (viz., the uncertainty in the abscissa of the electron) then one must have:

$$\Delta x \ll y, \tag{1}$$

and similarly:

$$\Delta y \ll y. \tag{1'}$$

Having said that, upon passing close to the coordinate origin, the electron must produce two effects on the magnetometer: One of them is due to the magnetic field that is created by the translatory motion of the electron (viz., the orbital field H_0); its value is:

$$H_0 = \frac{ev_x}{cy^2}. \tag{2}$$

The second effect is due to the proper magnetic moment of the electron that is created by the proper magnetic field at the place that is occupied by the magnetometer:

$$H_P = \frac{eh}{4\pi m_0 c} \cdot \frac{1}{y^3}. \tag{3}$$

However, the value of H_0 is not known with absolute precision. That quantity is governed by an uncertainty of:

$$\Delta H_0 = \frac{e}{c} \left| \frac{\Delta v_x}{y^2} + \frac{v_x}{y^3} \Delta y \right|, \tag{4}$$

since there is necessarily an uncertainty Δv_x in the x component of the velocity of the electron and an uncertainty Δy in its ordinate y ([1]). However, from the Heisenberg relations, if we limit ourselves to the Newtonian approximation then we will have:

$$\Delta v_x \geq \frac{h}{m_0 \Delta x}, \qquad \Delta v_y \geq \frac{h}{m_0 \Delta y}, \tag{5}$$

and on the other hand, we must always suppose that:

$$\Delta v_y \leq v_x, \tag{6}$$

since otherwise we would not have any appreciable motion along the x-axis.

If we now compare (3) and (4) then we will find that:

$$\frac{[\Delta H_0]}{H_P} = \frac{4\pi m_0}{h} (y \, \Delta v_x + v_x \, \Delta y), \tag{7}$$

so it will result from (5), (1), and (6) that:

$$\frac{[\Delta H_0]}{H_P} \geq \left(\frac{y}{\Delta x} + \frac{v_x}{\Delta v_y} \right) \gg 1. \tag{8}$$

The uncertainty in the orbital magnetic field is therefore always much larger than the value of the proper field, and in turn, the magnetometer will not permit us to measure the proper magnetic moment of the electron.

We repeat the same kind of argument for the behavior of the electron-magnet in a non-uniform magnetic field. Once again, let an electron be displaced along the x-axis. Imagine a magnetic field that is parallel to the y-axis and possesses a noticeable gradient $\partial H / \partial y$.

The magnetic field has the value $H(O)$ at the coordinate origin. Upon passing close to the origin, the moving electron will be subjected to the Lorentz electrodynamical force that is due to the orbital motion; that force will be reasonably equal to:

([1]) One must have the + sign inside the bracket in the formula (4), since the two uncertainties can combine their effects in the must unfavorable case.

$$f_0 = e \, \frac{v_x}{c} \cdot H(O). \tag{9}$$

On the other hand, the magnet to which the electron is equivalent is subject to the force (which is due to the proper magnetic moment):

$$f_P = \frac{eh}{4\pi m_0 c} \left(\frac{\partial H}{\partial y} \right)_{y=0}. \tag{10}$$

However, since the wave train has a size of Δy and the ordinate of the electron is, in turn, affected with that uncertainty, one will have the following uncertainty in the value of f_0:

$$| \Delta f_0 | = \frac{e}{c} \left| v_x \left(\frac{\partial H}{\partial y} \right)_0 \Delta y + H(O) \, \Delta v_x \right|, \tag{11}$$

so, upon comparing this with (10), one will get:

$$\frac{|\Delta f_0|}{f_P} = \frac{4\pi m_0}{h} \left| v_x \Delta y + \left(\frac{H}{\frac{\partial H}{\partial y}} \right)_0 \Delta v_x \right|. \tag{12}$$

Since the relations (5) and (6) are again valid here, one will then have:

$$\frac{|\Delta f_0|}{f_P} \geq \left| \frac{v_x}{\Delta v_y} + \left(\frac{H}{\frac{\partial H}{\partial y}} \right)_0 \cdot \frac{1}{\Delta x} \right| \gg 1. \tag{13}$$

The effect of the non-uniform, external, magnetic field on the electronic magnet is then masked completely by the uncertainty on the Lorentz force, and that method will no longer permit us to measure the proper magnetic moment.

These arguments, which one can, moreover, perfect ([1]), thus make it very likely that one cannot measure the proper magnetic moment of the electron directly.


2. Impossibility of measuring the proper angular momentum. – As Darwin has pointed out, one can extend the same kind of considerations to the proper moment of rotation.

([1]) See DARWIN: "Examples of the Uncertainty Principle," Proc. of the Royal Society A **130** (1931), pp. 637, and the report by Pauli to the Solvay Congress in 1930 (Gauthier-Villars, editor).

For example, imagine a rectangular opening that is cut out of a planar screen and whose edges are Δx and Δy.

Make an electron that is associated with a monochromatic plane wave fall upon the left edge; the wave train of the electron on the right of the screen will have lateral dimensions Δx and Δy. The components x and y of the quantity of motion of the electron after passing through the screen will thus be affected with uncertainties:

$$| \Delta p_x | \geq \frac{h}{\Delta x} \,, \qquad\qquad | \Delta p_y | \geq \frac{h}{\Delta y} \,. \tag{14}$$
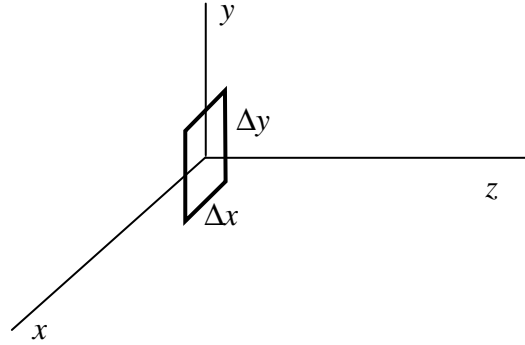
Now, the z component of the orbital angular momentum of the electron – i.e., the angular momentum that is due to its translatory motion – is:

$$M_z = x \, p_y - y \, p_x \,. \tag{15}$$

It will then have a value that is between 0 and:

$$\Delta M_z = \Delta x \, | \Delta p_y | + \Delta y \, | \Delta p_x | \,, \tag{16}$$

so, from (14):

$$\Delta M_z \geq h \left(\frac{\Delta x}{\Delta y} + \frac{\Delta y}{\Delta x} \right) = h \cdot \frac{\Delta x^2 + \Delta y^2}{\Delta x \, \Delta y} \,. \tag{17}$$

Since the fraction that appears in the right-hand side of (17) is greater than 1, one will deduce, *a fortiori*, that:

$$\Delta M_z > \frac{h}{4\pi} \,. \tag{18}$$

The uncertainty in the z component of the orbital angular momentum is therefore greater than the proper angular momentum, and that will make the measurement of that quantity illusory.

The same argument can be applied to the x and y components of the proper angular momentum.

These considerations seem to establish that the proper angular momentum of the electron is no more measurable experimentally than its proper magnetic moment.

3. More general theory. – As Pauli, in particular, has shown ([1]), one can apply the preceding considerations from a more general viewpoint. In order to understand it, we shall very briefly recall the method of approximation that is called the *Brillouin-Wentzel method*.

In non-relativistic wave mechanics, the principle of the Brillouin-Wentzel method consists of setting:

$$\Psi = e^{\frac{2\pi i}{h}S}, \qquad (19)$$

with:

$$S = S_0 + \frac{h}{2\pi i} S_1 + \left(\frac{h}{2\pi i}\right)^2 S_2 + \ldots + \left(\frac{h}{2\pi i}\right)^n S_n + \ldots \qquad (20)$$

and determining the terms S_0, S_1, … that figure in the development of S in successive powers of the quantity of very small modulus $h / 2\pi i$ by successive approximations. If one is content with terms of order zero in this development – i.e., the ones that persist when $h = 0$ – then one will recover geometrical optics for the associated corpuscle, since in that case one can consider a group of waves with very small dimensions to describe one of the rays of the wave and assimilate it with the point-like corpuscle of the old mechanics; the rays of the wave and the classical trajectories will coincide. However, if one takes into account the terms of order 1, 2, … in $h / 2\pi i$ in the development (20) then one will see some peculiarities appear that contrast wave optics with geometric optics and the new mechanics with the old.

One can repeat the same method of approximation in Dirac's theory; that is what Pauli did. Here, one sets:

$$\Psi_k = e^{\frac{2\pi i}{h}S_k} \qquad (k = 1, 2, 3, 4), \qquad (21)$$

with:

$$S_k = S_{0,k} + \frac{h}{2\pi i} S_{1,k} + \ldots + \left(\frac{h}{2\pi i}\right)^n S_{n,k} + \ldots, \qquad (22)$$

and one will determine the $S_{i,k}$ by successive approximations. Upon proceeding in that way, it seems that if one keeps only the terms of order 0 then one will obtain geometrical optics that corresponds to *relativistic* wave mechanics *without spin*; i.e., the old mechanics in the Einsteinian form. One can then imagine some extremely small wave trains that describe ray-trajectories that coincide with the trajectories that are predicted by the Einsteinian dynamics of the electron. Now, Einsteinian mechanics considers the electron to be a simple charged corpuscle and ignores its "spin." That could be predicted, since the proper magnetic moment and the proper angular momentum of the electron are proportional to h, and will disappear if one neglects the terms of order h.

If, while pursuing the approximation, one takes terms of order 1, 2, … in $h / 2\pi i$ into account in equation (22) then one will see some terms appear that translate into the existence of magnetism and the proper rotation of the electron, but at the same time, as in the non-relativistic theory, one will leave the domain in which geometrical optics applies, and as a consequence, the concepts of the old point-like mechanics will cease to be exact.

([1]) Helvetica Physica Acta, vol. V, fasc. III, pp. 179.

One now comprehends why, conforming to Bohr's conclusions, an experiment in which it is possible to treat the electron as a material point cannot lead to the exhibition of the proper magnetic moment or the proper rotation of the electron. Indeed, from Pauli's analysis, it would be a contradiction to assume that the old point-like mechanics is valid and that the characteristic effects of "spin" can be manifested.

4. Polarization of electronic waves. – If it seems impossible to measure the magnetic moment of an electron when it is considered to be a magnetic corpuscle then nothing, *a priori*, will oppose the experimental demonstration of the character of polarization that is imposed on the electronic wave Ψ by the existence of that magnetic moment. Moreover, the polarization of the wave Ψ differs noticeably from the classical polarization of light waves. Whereas the latter is defined for a plane wave by a vector that always oscillates normal to the direction of propagation, the polarization of the Dirac wave Ψ is defined by the vector **I** of its "magnetic moment density," and for a plane wave that vector will be, as we know, constant and oriented in no particular way with respect to the direction of propagation. As a consequence, whereas the properties of a sheaf of polarized light, when examined from the various azimuths around its direction of propagation, will always present the period π, the corresponding period for a sheaf of Dirac waves Ψ will be 2π.

In order to exhibit the polarization of electronic waves, one can contemplate the use of an arrangement that is analogous to the Nuremberg apparatus in optics. Suppose that one has a sheaf of unpolarized electrons – i.e., a sheaf in which the vectors **I** that relate to the various electrons are oriented at random. If one reflects that sheaf on a crystalline body then one can imagine that, by analogy with optics, the reflected sheaf can be partially polarized. If one makes the reflected sheaf fall upon a second reflector then the second reflection will happen with more or less intensity according to the azimuth of the plane of incidence. The exact theory of the phenomenon seems to be very complicated and has not been developed completely either; we shall not go into it here. Loosely speaking, one is led to predict a very weak effect of several millivolts for electrons, which is an effect that must increase with the energy. From the experimental viewpoint, the phenomenon does not seem to have been observed with any certainty. Rupp has indeed published some photographs in which one neatly sees the influence of the azimuth on the second reflection, but these results do not seem to have been confirmed up to now by those of other experimenters. The question remains open ([1]).

––––––––––––

([1]) The principal theoretical works on the polarization of electrons are those of Mott, Proc. Roy. Soc. A **124** (1924), pp. 425 and *ibid.*, A **135** (1932), pp. 429. One will find a complete bibliography in a memoir of Thibaud, Trillat, and v. Hirsch, J. Phys. Rad. (7) **3** (1932), pp. 314.

CHAPTER XX

NEGATIVE-ENERGY STATES IN DIRAC'S THEORY

1. Negative-energy plane waves

We shall now address one of the great problems that were suggested by Dirac's theory.

Previously, we studied the form of the functions Ψ_k for a monochromatic plane wave in the case of the absence of a field (see Chapter XII, paragraph 3). In order to do that, we wrote down the solution to the Dirac equation with zero potentials, and we tested a solution of the form:

$$\Psi_k = a_k \, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} . \tag{1}$$

We then found the following homogeneous, linear equations for the a_k :

$$\left(\frac{W}{c} + m_0 c\right) a_1 + (p_x + i\, p_y)\, a_4 + p_z\, a_3 = 0,$$

$$\left(\frac{W}{c} + m_0 c\right) a_2 + (p_x - i\, p_y)\, a_3 - p_z\, a_4 = 0,$$

$$\left(\frac{W}{c} + m_0 c\right) a_3 + (p_x + i\, p_y)\, a_2 + p_z\, a_1 = 0, \tag{2}$$

$$\left(\frac{W}{c} + m_0 c\right) a_4 + (p_x - i\, p_y)\, a_1 - p_z\, a_2 = 0.$$

In order for there to be a non-zero solution, it is necessary that the determinant of equations (2) must be zero. That will give us the condition:

$$\frac{W^2}{c^2} = m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2 , \tag{3}$$

which is the classical relativistic relation. Upon setting:

$$W = + c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2} , \tag{4}$$

we will then find the solution:

$$a_1 = -\frac{p_z A + (p_x + ip_y)B}{W/c + m_0 c}, \qquad a_2 = -\frac{(p_x - ip_y)A - p_z B}{W/c + m_0 c}, \qquad a_3 = A, \qquad a_4 = B, \quad (5)$$

where A and B are arbitrary complex constants.

However, we can also satisfy condition (3) by setting:

$$W = -c\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}. \tag{6}$$

We will then find the following solution:

$$a_1 = C, \qquad a_2 = D, \qquad a_3 = \frac{p_z C + (p_x + ip_y)D}{m_0 c - W/c}, \qquad a_4 = \frac{(p_x - ip_y)C - p_z D}{m_0 c - W/c}, \quad (7)$$

in which C and D are two arbitrary complex constants.

We shall now modify the notations that we just employed slightly. For given values of p_x, p_y, p_z, we shall henceforth always define W by formula (4) with the + sign, and in order to take the solution (7) into account, we say that we have to consider both the wave of energy + W and the wave of energy – W. With that new convention, one must change W into – W in equations (7).

Briefly, for a given value of p_x, p_y, p_z, while W is defined by relation (4), we must consider the monochromatic plane wave of positive energy + W that is defined by:

$$\Psi_k = a_k\, e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)}, \tag{8}$$

with:

$$a_1 = -\frac{p_z A + (p_x + ip_y)B}{W/c + m_0 c}, \qquad a_2 = -\frac{(p_x - ip_y)A - p_z B}{W/c + m_0 c}, \qquad a_3 = A, \qquad a_4 = B, \quad (9)$$

and the monochromatic plane wave with negative energy – W that is defined by:

$$\Psi_k = a_k\, e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)}, \tag{10}$$

with:

$$a_1 = C, \qquad a_2 = D, \qquad a_3 = \frac{p_z C + (p_x + ip_y)D}{W/c + m_0 c}, \qquad a_4 = \frac{(p_x - ip_y)C - p_z D}{W/c + m_0 c}. \quad (11)$$

We shall now examine these 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 positive proper mass + m_0, outweigh the waves Ψ_1 and Ψ_2, which correspond in some way to the negative proper mass – m_0 . These waves Ψ_1 and Ψ_2, which will be zero when the velocity is zero, will become important only for velocities that are sufficiently close to that of light. It will only be in the limiting case of $v = c$ that one will have:

$$| \Psi_1 |^2 + | \Psi_2 |^2 = | \Psi_3 |^2 + | \Psi_4 |^2.$$

In other words, one will always have:

$$\Omega_1 = | \Psi_1 |^2 + | \Psi_2 |^2 - | \Psi_3 |^2 - | \Psi_4 |^2 \leq 0, \qquad (12)$$

in which equality refers to the limiting case $v = c$.

The conclusions will be the opposite for the negative-energy wave. The waves Ψ_1 and Ψ_2 will dominate in that case. If the electron is at rest – i.e., if p_x, p_y, and p_z are zero – then one will have:

$$\Psi_1 = C \, e^{-\frac{2\pi i}{h} m_0 c^2 t}, \qquad \Psi_2 = D \, e^{-\frac{2\pi i}{h} m_0 c^2 t}, \qquad \Psi_3 = \Psi_4 = 0. \qquad (13)$$

The waves Ψ_3 and Ψ_4 become important for increasing velocities, but it will be only in the limiting case of $v = c$ that one will have equality between:

$$| \Psi_1 |^2 + | \Psi_2 |^2 \qquad \text{and} \qquad | \Psi_3 |^2 + | \Psi_4 |^2.$$

One will thus have:

$$\Omega_1 \geq 0 \qquad (14)$$

here, in which equality refers to the limiting case $v = c$. Here, it is the waves that correspond, in some way, to a negative proper mass that dominates.

The existence of negative-energy states in Dirac's theory constitutes a grave difficulty in that theory, since an electron that is in such a state will have properties that are completely foreign and have never been observed. When placed in an electric field **h**, it will take on an acceleration whose sense is opposite to the force – $e\mathbf{h}$, so one increases its velocity by reducing its energy, its velocity will be in the opposite sense to its quantity of motion ([1]), etc. It then seems that Dirac's theory forces one to eliminate these negative-energy states, since they do not answer to reality. However, for reasons that we shall now point out, that elimination does not seem easy.

2. The incomplete character of a system of positive-energy waves. – We return to certain peculiarities of non-relativistic wave mechanics.

In non-relativistic wave mechanics, the equation of propagation:

$$\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t} = H(\Psi) \qquad (15)$$

is of first order in time. The solution to that equation will thus be determined completely if one knows its initial form $\Psi(x, y, z, 0)$.

Consider the case of the absence of the field. Equation (15) will then be written:

([1]) By virtue of the relation $v = \partial W / \partial p$, which expresses the equality of the velocity of the corpuscle and the group velocity of the associated wave (Cf., *Introd. à l'etude de la Mécanique ondulatoire*, pp. 75).

$$\Delta\Psi = \frac{4\pi i\, m}{h}\frac{\partial\Psi}{\partial t} \tag{16}$$

and will admit the monochromatic plane wave:

$$\Psi(x,\, y,\, z,\, t) = a\; e^{\frac{2\pi i}{h}(Et - p_x x - p_y z - p_z z)} \tag{17}$$

for a solution with:

$$E = \frac{1}{2m}[p_x^2 + p_y^2 + p_z^2] \tag{18}$$

(in which, there is no sign ambiguity).

Suppose that we give the wave function the initial form:

$$\Psi(x,\, y,\, z,\, 0) = F(x,\, y,\, z). \tag{19}$$

Suppose that $F(x,\, y,\, z)$ is developable into a Fourier integral of the form:

$$F(x,\, y,\, z) = \int\!\!\int\!\!\int_{-\infty}^{+\infty} g(p_x,\, p_y,\, p_z)\, e^{-\frac{2\pi i}{h}(p_x x + p_y z + p_z z)}\, dp_x dp_y dp_z. \tag{20}$$

If one starts with the given function F then one can calculate the coefficients of this development by the formula:

$$g(p_x,\, p_y,\, p_z) = \frac{1}{h^3}\int\!\!\int\!\!\int_{-\infty}^{+\infty} F(x,\, y,\, z)\, e^{\frac{2\pi i}{h}(p_x x + p_y z + p_z z)}\, dx\, dy\, dz. \tag{21}$$

I then say that the function:

$$\Psi(x,\, y,\, z,\, t) = \int\!\!\int\!\!\int_{-\infty}^{+\infty} g(p_x,\, p_y,\, p_z)\, e^{\frac{2\pi i}{h}(Et - p_x x - p_y z - p_z z)}\, dp_x\, dy_y\, dz_z. \tag{22}$$

This should be obvious, since:

1. Ψ is a solution of equation (16), since it is a sum of monochromatic plane wave solutions of that linear equation.
2. For $t = 0$, one will have, in fact, $\Psi(x,\, y,\, z,\, 0) = F(x,\, y,\, z)$.

We conclude the following theorem from that:

In the non-relativistic wave mechanics of the free, material point, monochromatic plane waves will constitute a "complete" system; i.e., it will be possible to represent any solution by a superposition of such waves.

We now pass on to Dirac's theory and look for the translation of the argument that we just made into that context. There, one will have four simultaneous first-order equations in the four Ψ_k. These four wave functions are thus determined completely if one is given the initial form $\Psi_k(x, y, z, 0)$.

We always take the case of the zero field and further demand that it must be possible to represent any sort of solution by a superposition of monochromatic plane waves. An arbitrary solution is defined completely by the four functions:

$$\Psi_k(x, y, z, 0) = F_k(x, y, z) \qquad\qquad (k = 1, 2, 3, 4). \qquad\qquad (23)$$

Suppose that the F_k are given arbitrarily and developed by Fourier's theorem into the form:

$$F_k(x, y, z) = \int\int\int_{-\infty}^{+\infty} g_k(p_x, p_y, p_z)\, e^{-\frac{2\pi i}{h}(p_x x + p_y z + p_z z)}\, dp_x\, dp_y\, dp_z, \qquad\qquad (24)$$

in which the g_k are given by:

$$g_k(p_x, p_y, p_z) = \frac{1}{h^3}\int\int\int_{-\infty}^{+\infty} F_k(x, y, z)\, e^{\frac{2\pi i}{h}(p_x x + p_y z + p_z z)}\, dx\, dy\, dz. \qquad\qquad (25)$$

We seek to represent the solution that corresponds to the given initial Ψ_k by a superposition of monochromatic plane waves that contains only positive-energy waves. In order to this, we must have:

$$\Psi_k(x, y, z, t) = \int\int\int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y z - p_z z)}\, dp_x\, dy_y\, dz_z, \qquad\qquad (26)$$

in which W is defined by relation (4). These functions (26) indeed give a solution to the equation of propagation, but in order for it to take the initial form $F_k(x, y, z)$, one must have:

$$a_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z) \qquad\qquad (k = 1, 2, 3, 4), \qquad\qquad (27)$$

in which the g_k are known. However, we know that only two of the four $a_k(p_x, p_y, p_z)$ be chosen arbitrarily, and that will show us that one cannot satisfy the conditions (27), in general. The positive-energy, monochromatic, plane waves do not form a complete system for the Dirac electron in the absence of a field then.

On the contrary, we will obtain a complete system if we consider the negative-energy plane waves, as well as the positive-energy ones. Indeed, if we set:

$$\Psi_k(x, y, z, t) = \int\int\int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y z - p_z z)}\, dp_x\, dy_y\, dz_z$$

$$+ \int\int\int_{-\infty}^{+\infty} b_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y z - p_z z)}\, dp_x\, dy_y\, dz_z, \qquad\qquad (28)$$

then, since W is always defined by (4), we will get a solution to the equations of propagation, since they are linear, and in order for the initial Ψ_k to coincide with the given $F_k(x, y, z)$, we must write the conditions:

$$a_k(p_x, p_y, p_z) + b_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z) \qquad (k = 1, 2, 3, 4). \qquad (29)$$

Now, conditions (29), unlike conditions (27), are compatible, because four of the eight a_k and b_k that correspond to a set of p_x, p_y, p_z are arbitrary.

If we write the conditions (29) explicitly then we will obtain the following conditions for every set of p_x, p_y, p_z :

$$
\begin{aligned}
-\frac{(p_x + ip_y)A + p_z B}{W/c + m_0 c} + C &= g_1(p_x, p_y, p_z), \\[2em]
-\frac{(p_x - ip_y)A - p_z B}{W/c + m_0 c} + D &= g_2(p_x, p_y, p_z), \\[2em]
\frac{p_z C + (p_x + ip_y)D}{W/c + m_0 c} + A &= g_3(p_x, p_y, p_z), \\[2em]
\frac{(p_x - ip_y)C + p_z C}{W/c + m_0 c} + B &= g_4(p_x, p_y, p_z).
\end{aligned}
\qquad (30)
$$

If one studies the system (30), while taking the Heisenberg relations into account, then one will see that in the most favorable cases it will be possible to represent a wave train Ψ by a superposition of positive-energy waves, but only if the dimensions of the wave train are appreciably larger than $h / m_0 c$. By contrast, if the dimensions of the wave train are less than $h / m_0 c$ then it will generally be entirely impossible to represent it by a superposition of plane waves without appealing to negative-energy waves.

It is therefore impossible in Dirac's theory to represent an arbitrary wave train in a general fashion without involving negative-energy waves, and that impossibility shows us why it is difficult to eliminate those waves.

3. The Klein paradox. – In the first place, one must recognize that the problem of negative-energies exists even in classical relativity. Indeed, in classical relativity, energy is defined as a function of the quantity of motion in the case of the absence of the field by the relation (3), which is a relation that gives the following two values to W:

$$W = \pm c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}. \qquad (31)$$

However, one must remark that, from (31), the possible values of W will be comprised of two separate domains, namely, from $+ \infty$ to $+ m_0 c^2$ and from $- m_0 c^2$ to $-$

∞. The interval from $- m_0 \, c^2$ to $+ m_0 \, c^2$ does not correspond to any possible value of energy. Now, in the old mechanics – even relativistically – the mechanical quantities, and in particular, the energy, vary in a continuous fashion, in principle. Therefore, if the electrons have energies at the origin that are found in the positive domain $m_0 \, c^2 \to + \infty$ then the same thing will always be true in what follows, and no value of the energy that is found in the negative domain $- m_0 \, c^2 \to - \infty$ can appear, since the two domains are separated by the interval $m_0 \, c^2 \to - m_0 \, c^2$, which the values of energy cannot cross. The objection that was raised against the existence of negative energies is thus found to be lifted in Einsteinian dynamics.

The same thing is not true in the new mechanics, because, in principle, it admits the possibility of brief transitions between states whose energy differs by a finite quantity, which obstructs the *a priori* elimination of the passage from the domain of positive energies to the one of negative energies. Moreover, it is easy to imagine some simple examples in which transitions of that genre are found to be realized.

O. Klein was the first ([1]) to point out an example of that transition, which, without being a transition from a positive-energy state to a negative-energy state, properly speaking, was nonetheless equivalent to one.

Klein considered a planar surface S that separated a region I in which the potential is zero from a region II in which one finds a constant scalar potential V that is also negative in such a way that an electron will have a potential energy of $U = - eV > 0$ in region II. A Dirac electronic wave that comes from region I falls normally upon the separation surface; that wave is assumed to be monochromatic and planar and to correspond to a positive energy of W. One then calculates the waves that are reflected and transmitted by the separation surface. One shows that in order to carry out the calculation, it is first necessary to express the idea that there is continuity for each of the four Ψ_k upon crossing the separation surface; i.e., to write the four equations:

$$\Psi_k \, (\text{incident}) + \Psi_k \, (\text{reflected}) = \Psi_k \, (\text{transmitted}). \tag{32}$$

Naturally, the reflected and transmitted waves correspond to the same energy as the incident wave; i.e., the phenomenon is conservative.

Having posed that, Klein proved the following results, which I shall be content to state without proof: For $0 < U < W - m_0 \, c^2$, there is both reflection and transmission, and the transmitted wave, like the reflected wave, will have the usual character of a positive-energy wave.

For $W - m_0 \, c^2 < U < W + m_0 \, c^2$, there will then be total reflection with a vanishing wave in the second medium.

For $U > W + m_0 \, c^2$, one will once more find a transmitted wave that crosses the surface, but – and this is the fundamental result – that wave will correspond to the total energy W, which is positive, but which one can call "the energy of a non-potential nature" of the electron in region II; i.e., the quantity $W - U$ will be negative and less than $- m_0 \, c^2$, whereas in classical Einsteinian dynamics, that quantity would always be greater than $m_0 c^2$. The transmitted wave in the medium II, which is governed by the scalar potential V, is analogous to a negative-energy wave in the absence of a potential, and will

---

([1])   Zeitschrift für Physik **53** (1929), pp. 157.

possess the same paradoxical properties. The existence of that transmitted wave must be interpreted by saying that there is a certain probability for an incident electron to penetrate into the region II by passing through that strange state,

For Klein, that probability can itself be noticeable. Indeed, it seems that one cannot consider the result of the calculation to be physically exact, and that is the Klein paradox.

It is true that the case that Klein considered was extremely schematic. Other authors imagined other examples that were a little less artificial. The general result of that research seems to be the following: Whenever the potential energy of the electron is subjected to a variation of at least $m_0 c^2$ over a distance of less than $h / m_0 c$, there will be the possibility that negative-energy states will appear. That result leads one to think that if one can forbid the consideration of spatial distances that are less than $h / m_0 c$ then one might succeed in eliminating these negative-energy waves. This must be compared to what was said at the end of the preceding paragraph on the subject of the representation of wave trains.

4. Remarks and conclusions. – The negative-energy states appear once more in a curious fashion in the theory of diffusion of light by the Dirac electron. We shall not develop that theory, but refer the reader to some other papers ([1]). We note only the result: The Dirac electron cannot diffuse light unless it is capable of taking on negative-energy states. Since it is necessary to assume that electrons diffuse light in order to explain the phenomenon of diffusion by material bodies, that fact by itself will once more show how difficult it is to liberate Dirac's theory from the apparent imperfection that takes the form of the existence of negative-energy states. Various attempts have nevertheless been made to get around the problem. We shall say only a few words about them.

Schrödinger proposed a very ingenious modification of the general Dirac equations that made the negative-energy states disappear ([2]). However, in addition to the fact that his modification seemed difficult to reconcile with the existence of the diffusion of light by electrons, it had a very neatly artificial character.

Dirac, instead of wishing to suppress the negative-energy states, sought to interpret them ([3]). In order to do that, he supposed that these states actually existed, and that in every subset of space one will find an infinite number of electrons that occupy all of these negative-energy states, which will be unobservable electrons. Some of these electrons leave their usual negative energy states from time to time in order to take on a positive-energy state, and those will be observable electrons. The "hole" that is left in the negative-energy state by the departure of an electron will be what one calls a proton [†]. The return of the missing electron to a negative-energy state will constitute the simultaneous disappearance of an electron and a proton, which must then be

---

([1]) Notably, see the remarkable treatise of E. Fermi: "Quantum theory of radiation," Reviews of Modern Physics **4** (1932), pp. 120.

([2]) Notably, see Annales de l'Institut Henri Poincaré, t. II, pp. 269.

([3]) *Ibidem*, t. I, pp. 357.

[†] Translator's note: This was not my mis-translation of the word "positron," but a common misunderstanding during the era between the publication of Dirac's first papers on the theory of the electron and the discovery of the positron. (See the following comment by de Broglie.)

accompanied by the emission of radiation. Unfortunately, these seductive hypotheses are riddled with all sorts of objections and do not seem to be capable of being saved ([1]).

In summary, the negative-energy states of the electron play a very important role in the structure itself of Dirac's theory, although they do not seem to manifest themselves in reality. It seems that, for the moment, one cannot pose the problem without solving it.

_____

_____

([1]) Dirac's theory of holes is enjoying a return to favor at the moment as a result of the experimental discovery of the positive electron.

SCHRÖDINGER'S ZITTERBEWEGUNG

———

1. The motion of the center of gravity of the probability.

In this chapter, we shall avoid explicitly introducing an operator that would correspond to the "velocity" of the electron. As Bohr has remarked, the velocity of a corpuscle is a notion that one must use prudently in the new mechanics. Indeed, it is well-defined only in certain cases, and it does not seem justified to consider it to be an observable physical quantity.

On the contrary, it is always permissible to imagine the mean position of a corpuscle – or the center of gravity of the probability – and to study its motion. In effect, that point is defined by its coordinates:

$$\overline{x} = \int\int_{-\infty}^{+\infty}\int x\sum_{k=1}^{4}\Psi_k^*\Psi_k\, d\tau\,, \qquad \overline{y} = \int\int_{-\infty}^{+\infty}\int y\sum_{k=1}^{4}\Psi_k^*\Psi_k\, d\tau\,, \qquad \overline{z} = \int\int_{-\infty}^{+\infty}\int z\sum_{k=1}^{4}\Psi_k^*\Psi_k\, d\tau \qquad (1)$$

in Dirac's theory.

The velocity of the point with coordinates $\overline{x}$, $\overline{y}$, $\overline{z}$ is itself a perfectly well-defined quantity.

Let $(\Psi_1, \Psi_2, \Psi_3, \Psi_4)$ be the solution to the Dirac equation that represents the wave that is associated with the motion of a certain electron. As we know, we can develop each Ψ_k in a series of proper functions of the form:

$$\Psi_k = \sum_{n} c_n \Psi_{k,n}\,, \qquad (2)$$

where the c_n are complex constants. Upon substituting this in (1), one will get:

$$\overline{x} = \sum_{m,n} c_m^* c_n \int_D \sum_{k=1}^{4} (\Psi_{k,m}^* x\, \Psi_{k,n})\, d\tau\,, \qquad (3)$$

in which we have replaced the triple integral of the summation in formulas (1) with just one symbol. Upon denoting the element of the matrix that corresponds to x and has the indices m, n by x_{mn}, formula (3) will be written simply as:

$$\overline{x} = \sum_{m,n} c_m^* c_n\, x_{mn}\,. \qquad (4)$$

Now, by virtue of formula (25) in Chapter XV, one will have:

$$\frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^{4} \left[\Psi_{k,m}^* \frac{2\pi i}{h}(xH - Hx)\Psi_{k,n} \right] d\tau , \qquad (5)$$

where H is Dirac's Hamiltonian operator:

$$H = - \left[e\, V + c\, (\alpha_1\, P_1 + \alpha_2\, P_2 + \alpha_3\, P_3 + \alpha_4\, m_0\, c) \right] . \qquad (6)$$

One easily finds:

$$\frac{2\pi i}{h}(xH - Hx) = c\, \alpha_1 \left(x \cdot \frac{\partial}{\partial x} - \frac{\partial}{\partial x}x \right) = -\, c\, \alpha_1 , \qquad (7)$$

and in turn:

$$\frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^{4} \left[\Psi_{k,m}^*(-c\,\alpha_1)\Psi_{k,n} \right] d\tau . \qquad (8)$$

On then obtains from (4):

$$\frac{d\overline{x}}{dt} = \sum_{m,n} c_m^*\, c_n \frac{dx_{mn}}{dt} = \int_D \sum_{k=1}^{4} \left[c_m^*\, \Psi_{k,m}^*(-c\,\alpha_1)\sum_n c_n\Psi_{k,n} \right] d\tau$$

$$= \int_D \sum_{k=1}^{4} \Psi_{k,m}^*(-c\,\alpha_1)\Psi_{k,n}\, d\tau . \qquad (9)$$

Naturally, one likewise finds that:

$$\frac{d\overline{y}}{dt} = \int_D \sum_{k=1}^{4} \Psi_k^*(-c\,\alpha_2)\Psi_k\, d\tau , \qquad \frac{d\overline{z}}{dt} = \int_D \sum_{k=1}^{4} \Psi_k^*(-c\,\alpha_3)\Psi_k\, d\tau . \qquad (10)$$

We previously found [formulas (7) of Chap. XII] the following expressions for the components of the probability current:

$$\rho\, u_x = -\, c \sum_{k=1}^{4} \Psi_{k,m}^*\alpha_1\Psi_{k,n}\, d\tau . \qquad (11)$$

Upon comparing this with (9), one will then get:

$$\frac{d\overline{x}}{dt} = \int_D \rho\, u_x\, d\tau = \overline{u}_x , \qquad (12)$$

and similarly:

$$\frac{d\overline{y}}{dt} = \overline{u}_y , \qquad \frac{d\overline{z}}{dt} = \overline{u}_z . \qquad (12')$$

The velocity of the center of gravity of the probability is then equal to the mean value of the velocity of the probability, which is a result that one could have predicted, *a priori*.

One often interprets formulas (9) and (10) 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 electron. Since these operators have only $+ c$ and $- c$ for proper values, one is led to say that from the general principles of new mechanics, the only possible values for the components of velocity will be $+ c$ and $- c$, which is a result that is very difficult to comprehend. As we have said, we prefer to abstain from making the operator correspond to the "velocity" of a corpuscle.

2. Ehrenfest's theorem is no longer exact in Dirac's theory. – In non-relativistic wave mechanics, we proved Ehrenfest's theorem, which is expressed by the formulas:

$$m\,\frac{d^2\overline{x}}{dt^2} = \overline{f}_x, \qquad m\,\frac{d^2\overline{y}}{dt^2} = \overline{f}_y, \qquad m\,\frac{d^2\overline{z}}{dt^2} = \overline{f}_z. \qquad (13)$$

If we apply this to the case in which a field is absent then this theorem will lead us to the following result: *In the absence of a field, the motion of the center of gravity is uniform and rectilinear.* This is, in a sense, the translation of the principle of inertia into wave mechanics.

The preceding result is no longer exact in Dirac's theory, in general. Indeed, start with formula (8) and once more apply formula (25) of Chapter XV; we will get:

$$\frac{d^2 x_{m,n}}{dt^2} = \int_D \sum_{k=1}^4 \left[\Psi_{k,m}^* \cdot \frac{2\pi i}{h}(- c\,\alpha_1 H + Hc\alpha_1) \cdot \Psi_{k,n} \right] \cdot d\tau, \qquad (14)$$

and, in turn, from (4):

$$\frac{d^2 x}{dt^2} = \sum_{m,n} c_m^* c_n \int_D \sum_{k=1}^4 \left[\Psi_{k,m}^* \frac{2\pi i}{h}(- c\,\alpha_1 H + Hc\alpha_1) \Psi_{k,n} \right] \cdot d\tau. \qquad (15)$$

Now, α_1 does not commute with H, since it anti-commutes with α_2, α_3, and α_4. One thus has, in general:

$$\frac{d^2\overline{x}}{dt^2} \neq 0, \qquad \text{and similarly:} \quad \frac{d^2\overline{y}}{dt^2} \neq 0, \qquad \frac{d^2\overline{z}}{dt^2} \neq 0. \qquad (16)$$

The motion of the center of gravity of the probability in the absence of a field is not generally uniform and rectilinear.

We shall now examine more closely the reason that prevents this motion of the center of gravity from being uniform and rectilinear, and we shall see that it is related to the existence of negative-energy states. We will thus obtain some results that were obtained by Schrödinger in a very different form [1].

---

[1] Annales de l'Institut H. Poincaré, *loc. cit.*

3. Schrödinger's zitterbewegung. – In order to better comprehend why the center of gravity of the probability does not generally have a uniform, rectilinear motion in Dirac mechanics – even in the absence of a field – we shall subject the expression:

$$\frac{d\overline{x}}{dt} = \int_D \sum_{k=1}^{4} \Psi_k^* (-c\,\alpha_1) \Psi_k \, d\tau \tag{17}$$

to a detailed analysis.

We know that one can always write:

$$\Psi_k(x, y, z, t) = \int\int\limits_{-\infty}^{+\infty}\int \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}[Wt - p_x x - p_y y - p_z z]} \right.$$

$$\left. + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}[Wt - p_x x - p_y y - p_z z]} \right] dp_x \, dp_y \, dp_z \,, \tag{18}$$

with

$$W = + \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2} \,. \tag{19}$$

The eight $a_k (p_x, p_y, p_z)$ and $b_k (p_x, p_y, p_z)$ can be calculated by starting with four of them that are arbitrary using formulas that we already know.

Now call the space that is defined by the rectangular coordinates of the quantities p_x, p_y, p_z "momentum space," and divide that space into cells σ that are as small as we desire. The quantities:

$$\Delta(\sigma) = \iiint e^{-\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dp_x \, dp_y \, dp_z \tag{20}$$

are (up to a normalization constant) the "proper differentials" of the continuous spectrum of the monochromatic plane waves ([1]), and we can write:

$$\Psi_k (x, y, z, t) = \sum_k \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}Wt} \right] \Delta(\sigma) \,, \tag{21}$$

in which p_x, p_y, p_z are the coordinates of the center of the element σ in moment space, and $\sum\limits_{\sigma}$ denotes the summation over all cells σ in that space.

If that is true then we can write (17) in the form:

$$\frac{d\overline{x}}{dt} = -c \sum_\sigma \sum_{\sigma'} \sum_{k=1}^{4} \left[a_k(p_x', p_y', p_z') e^{\frac{2\pi i}{h}Wt} + b_k(p_x', p_y', p_z') e^{-\frac{2\pi i}{h}Wt} \right]^*$$

$$\cdot \alpha_1 \sum_k \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}Wt} \right] \cdot \int_D \Delta^*(\sigma') \Delta(\sigma) d\tau \,, \tag{22}$$

([1])   See the definition of the proper differential in Chapter V, paragraph 4.

where the domain is naturally all of space here. Since the proper differentials are orthogonal and can be assumed to be normalized, we have ([1]):

$$\frac{d\overline{x}}{dt} = -c \sum_\sigma \sigma \sum_{k=1}^{4} \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}Wt} \right]^*$$

$$\cdot \alpha_1 \left[a_k(p_x, p_y, p_z) e^{\frac{2\pi i}{h}Wt} + b_k(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}Wt} \right], \tag{23}$$

or:

$$\frac{d\overline{x}}{dt} = -c \sum_\sigma \sigma \sum_{k=1}^{4} \left[a_k^* \alpha_1 a_k + b_k^* \alpha_1 b_k \right] - c \sum_\sigma \sigma \left(\sum_{k=1}^{4} a_k^* \alpha_1 a_k e^{-\frac{4\pi i}{h}Wt} + \sum_{k=1}^{4} b_k^* \alpha_1 b_k e^{\frac{4\pi i}{h}Wt} \right). \tag{24}$$

We must now transform this expression. From the formulas of the preceding chapter, we have:

$$a_1 = -\frac{p_z A + (p_x - ip_y)B}{W/c + m_0 c}, \qquad a_2 = -\frac{(p_x - ip_y)A - p_z B}{W/c + m_0 c},$$

$$a_3 = A, \qquad a_4 = B, \tag{25}$$

so:

$$-c \sum_{k=1}^{4} a_k^* \alpha_1 a_k = -c \, (a_1^* a_4 + a_2^* a_3 + a_3^* a_2 + a_4^* a_1) = 2 p_x c^2 \frac{AA^* + BB^*}{W + m_0 c^2}. \tag{26}$$

However, one also has:

$$\sum_{k=1}^{4} a_k^* a_k = (AA^* + BB^*) \left(1 + \frac{p_x^2 + p_y^2 + p_z^2}{(W/c + m_0 c)^2} \right) = 2W \frac{AA^* + BB^*}{W + m_0 c^2}, \tag{27}$$

so, upon comparing this with (26), one will get:

$$-c \sum_{k=1}^{4} a_k^* \alpha_1 a_k = \frac{p_x c^2}{W} \sum_{k=1}^{4} a_k^* a_k . \tag{28}$$

One likewise finds upon using the expression for b_k that:

$$-c \sum_{k=1}^{4} b_k^* \alpha_1 b_k = -\frac{p_x c^2}{W} \sum_{k=1}^{4} b_k^* b_k . \tag{29}$$

On the other hand, the last two terms in the expression (24) are complex conjugates (due to the Hermiticity of α_1), and one can write:

([1]) Here, we denote the volume of the cell σ by σ.

$$c \sum_{\sigma} \sigma A_1 \cos\left(\frac{4\pi}{h} Wt + \varphi_1 \right). \tag{30}$$

Naturally, A_1 and φ_1 vary from one cell to another; i.e., they are functions of p_x, p_y, p_z. Finally, from (28), (29), and (30), we can write (24) in the form:

$$\frac{dx}{dt} = \sum_{\sigma} \sigma \frac{c^2 p_x}{W} \left(\sum_{k=1}^{4} a_k^* a_k - \sum_{k=1}^{4} b_k^* b_k \right) + \sum_{\sigma} \sigma c A_1 \left(\frac{4\pi W}{h} t + \varphi_1 \right). \tag{31}$$

Now, from the formulas of relativistic dynamics, the quantity $c^2 p_x / W$ is the x component of the velocity that corresponds to the quantity of motion p_x and to an energy $+ W$. Likewise, one can consider $- c^2 p_x / W$ to be the x component of a velocity that corresponds to the same quantity of motion and a negative energy $- W$. The first term in the expression (31) for dx/dt is therefore a sort of mean value for the component of the velocity v_x that corresponds to the spectral decomposition (18) and the wave Ψ. We then set:

$$\overline{v}_x = \sum_{\sigma} \sigma \frac{p_x c^2}{W} \sum_{k=1}^{4} (a_k^* a_k - b_k^* b_k) ; \tag{32}$$

naturally, $\overline{v}_x$ is independent of time. We then have:

$$\frac{dx}{dt} = \overline{v}_x + \sum_{\sigma} \sigma c A_1 \cos\left(\frac{4\pi}{h} Wt + \varphi_1 \right), \tag{33}$$

so, by integration:

$$\overline{x} = \text{const.} + \overline{v}_x t + \sum_{\sigma} \sigma \frac{hc A_1}{4\pi W} \sin\left(\frac{4\pi}{h} Wt + \varphi_1 \right). \tag{34}$$

Set:

$$x = \text{const.} + \overline{v}_x t . \tag{35}$$

One gets:

$$\overline{x} = \xi + \sum_{\sigma} \sigma \frac{hc A_1}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} \cdot t + \varphi_1 \right]. \tag{36}$$

One will likewise find:

$$\overline{y} = \eta + \sum_{\sigma} \sigma \frac{hc A_2}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} \cdot t + \varphi_2 \right],$$

$$\overline{z} = \zeta + \sum_{\sigma} \sigma \frac{hc A_3}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} \cdot t + \varphi_3 \right], \tag{37}$$

with the definitions:

$$\eta = \overline{v}_y t + \text{const.}, \qquad \overline{v}_y = \sum_{\sigma} \sigma \frac{p_y c^2}{W} \cdot \sum_{k=1}^{k} (a_k^* a_k - b_k^* b_k), \tag{38}$$

$$\zeta = \overline{v}_z t + \text{const.}, \qquad \overline{v}_z = \sum_\sigma \sigma \frac{p_z c^2}{W} \cdot \sum_{k=1}^{k} (a_k^* a_k - b_k^* b_k).$$

The point whose coordinates are ξ, η, ζ is displaced with a uniform, rectilinear motion, but the center of gravity of the probability will exhibit a series of oscillations about that point with a frequency of $2W / h$: This is "Schrödinger's *zitterbewegung*." Moreover, the amplitudes of these oscillations are weak, in general, because they will be proportional to the factor $hc / 4\pi W$, which is always smaller than $\dfrac{hc}{4\pi m_0 c} = \dfrac{1}{4\pi} \cdot \dfrac{h}{m_0 c}$.

Now, the quantity $h / m_0 c$ is often called the "Compton wave length" and is very small $(2.4 \cdot 10^{-12}$ cm.$)$.

The preceding analysis shows the origin of Schrödinger's zitterbewegung very neatly: It is due to the *beating* of the waves of positive energy W with the corresponding waves of negative energy $- W$. The beat frequency is, as usual, the difference of the frequencies of the beating waves, which is $2W / h$, here.

There is no Schrödinger zitterbewegung for a train of waves in a spectral decomposition that involves no negative-energy waves, and Ehrenfest's theorem will remain valid. However, we know that in the general case, one must appeal to negative-energy waves in order to represent a wave train. That is why the theorem is not generally valid in Dirac's theory.

Schrödinger's zitterbewegung and the non-validity of Ehrenfest's theorem are thus coupled to the existence of negative-energy states, and will disappear with those states if one can eliminate them.

_____

CHAPTER XXII

SOME REMARKS ON RELATIVIVY
AND THE NEW MECHANICS

———

1. The absence of symmetry between space and time in the new mechanics.

We have seen that Dirac's theory conforms to the principle of relativity in certain aspects. Indeed, one can give its fundamental equations a form that is invariant under a Lorentz transformation: One can form quantities from its four Ψ that have a tensorial character in space-time. Nevertheless, one can hardly pretend that Dirac's theory, in its present state, is in complete accord with the concepts of relativity, even in its special form. Indeed, one of the guiding ideas of relativity theory seems to be that one should always involve the coordinates of space and time in a symmetric fashion. Now, that symmetric intervention of the variables x, y, z, t is not realized in Dirac's theory, because one assumes the general principles of new mechanics, which are general principles that (at least, in their present form) give an entirely special role to the variable of "time." We must insist upon that point.

First, the new mechanics makes any observable physical quantity correspond to a Hermitian operator. Now, the Hermiticity of an operator is defined in a *spatial* domain, and that will already suffice to make the definition itself of some operators that are employed in the new mechanics non-relativistic. Time enters into those operators only as a parameter, and the derivatives $\partial^n / \partial t^n$ can never enter into them.

Having defined the operators that correspond to observable quantities, the new mechanics assumes that the possible values of one of those quantities are given by the proper values of the operator. However, the proper values and the proper functions of an operator are, in turn, defined in a domain D of space. The variable "time" plays no role in the calculation of proper values and proper functions of a Hermitian operator; when it does enter in, it enters only as a parameter. Once the possible values of an observable quantity A have been calculated, the new mechanics supposes that the various respective probabilities of the various possible values of that quantity will be given by the square of the modulus of the coefficient of each proper function in the development of the wave function Ψ in proper functions.

In the general case ([1]), those probabilities depend upon the parameter t, and it is precisely for that reason that the state of the system generally evolves.

Briefly: Up to now, the new mechanics has considered time to be something that plays a completely different role from that of the spatial coordinates. Dirac's theory, which assumes the general principles of new mechanics, cannot be truly "relativistic" then. One sees that very easily, for example, when one studies the mean values of quantities in Dirac's mechanics. The mean-value densities are defined in space, and in order to pass from the density to the mean value itself, one must integrate over space,

———

([1]) That is, when A is not a first integral.

which is an operation that is not relativistically invariant. One can then explain certain very obvious peculiarities in the table of mean values that is found at the end of Part II of this book, such as the very asymmetric use of the operator α_4, for example.

2. The fourth uncertainty relation. – We shall once more raise a question that relates to the particular role that is played by time in wave mechanics.

The Heisenberg uncertainty relations are the following ones:

$$\Delta p_x \cdot \Delta x \geq h, \quad \Delta p_y \cdot \Delta y \geq h, \quad \Delta p_z \cdot \Delta z \geq h . \tag{1}$$

From the mathematical standpoint, they are derived from the fact that wave mechanics makes the classical quantities p_x, p_y, p_z correspond (up to a constant) to the derivatives with respect to the conjugate variables x, y, z, respectively.

From the physical standpoint, the relations (1) must be interpreted by saying: At a given instant, the product of the uncertainty in one of the coordinates of a corpuscle with the conjugate component of the quantity of motion is always at least of order h.

The relativistic symmetry between space and time demands that the three relations (1) must be completed by a fourth relation:

$$\Delta W \cdot \Delta t \geq h, \tag{2}$$

since the energy W is the temporal component of a "world impulse" quadri-vector whose spatial components are p_x, p_y, p_z.

However, in the present state of the new mechanics, that fourth uncertainty relation cannot at all be interpreted in the same way as the first three, because, on the one hand, time t must be considered to be a parameter that has a well-defined value with no actual uncertainty, and on the other hand, the quantity "energy W" will correspond to the Hamiltonian operator, and not to the operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$, which cannot be considered to be Hermitian in the proper sense of the word.

It is nevertheless possible to give meaning to the equality in (2). Indeed, it is well-known that if one observes the passage of a wave from a fixed point in space during an arbitrary finite time interval Δt then one can assert only that the wave possesses a frequency v with an uncertainty of:

$$\Delta v \geq \frac{1}{\Delta t} . \tag{3}$$

For a wave Ψ, if one takes into account the relation $W = hv$ then one can write the inequality (3) in the form:

$$\Delta W \geq \frac{h}{\Delta t} . \tag{4}$$

One then perceives the sense of the relation (2). It expresses the idea that an experiment or observation that is made at a fixed point during an interval Δt cannot reveal the energy of a corpuscle with an uncertainty that is less than $h / \Delta t$.

The fourth uncertainty relation then has a meaning, but it is a very different meaning from that of the first three. That is a new aspect of the asymmetry between space and time in wave mechanics.

3. Can one establish symmetry between space and time in the new mechanics? – In order to establish symmetry between space and time in the new mechanics, one must try to modify the general principles in that sense. Without stating that this is impossible, we shall show that there are some great difficulties associated with that ambition.

The first thing to do will be to define the Hermiticity of operators in space-time, instead of defining it in space. For example, let an operator A act upon both the time and space variables (and also possibly upon the spin variable ζ). One will then have to define the Hermiticity by the condition:

$$\iiiint\limits_{D} f^{*}(x, y, z, t)\, A\, g\,(x, y, z, t)\, dx\, dy\, dz\, dt$$

$$= \iiiint\limits_{D} g\,(x, y, z, t)\, A^{*} f^{*}(x, y, z, t)\, dx\, dy\, dz\, dt. \tag{5}$$

D is a space-time domain, and for the free electron, it will be all of space-time.

One might possibly need to add a summation over the variable ζ in the condition (5).

The proper values and proper functions of the operator A will then be defined by the equation:

$$A\,[\varphi_i\,(x, y, z, t)] = \alpha_i\, \varphi_i\,(x, y, z, t), \tag{6}$$

in which φ_i is finite, uniform, continuous, and zero on the boundary of the space-time domain D. The wave function $\Psi\,(x, y, z, t)$ is developed in the form:

$$\Psi\,(x, y, z, t) = \sum_{i} c_i\, \varphi_i\,(x, y, z, t), \tag{7}$$

and one can transpose the usual statement of the general principles by saying:

A measurement of the physical quantity that corresponds to the operator A can give only one of the values α_i, and the probability of the value α_k will be equal to $|c_k|^2$.

It will result from this that the mean value of the quantity A will be:

$$\overline{A} = \iiiint\limits_{D} \Psi^{*} A\,(\Psi)\, dx\, dy\, dz\, dt\,. \tag{8}$$

Unfortunately, there is a grave objection to this: The quantities α_i and c_i, with their new definitions, are independent of time, and naturally the same thing will be true for $\overline{A}$. We will then get a theory of physical statics from which all time evolution is banned.

A different aspect of the difficulty is the following one: When we "quantize" a system with the current principles of the new mechanics – a hydrogen atom, for example – we mentally isolate that system from the rest of the universe. Rigorously speaking, that is not permissible: In order to determine Ψ, one must, in principle, take into account not only the force field that is created by the nucleus, but also all of the force fields that exist in the entire universe. Fortunately, the influence of the force fields that are external to the atom on the form of the stationary waves Ψ in the atom is completely negligible, because the waves Ψ tend to zero very rapidly when one extends them outside of the atomic domain. In principle, the determination of the stationary waves – viz., proper functions – demands the consideration of all space and everything that it contains, but in practice, the structure of the material world allows one to cut away systems that are sufficiently independent of the rest of the material world and consider them to be isolated. However, if we would like to define proper functions in space-time then that would no longer be true, because it does not seem possible to decompose the existence of a physical individual – such as an atom – into mutually-independent sections. Consider a hydrogen atom: In the course of its history, it will be subjected to various effects, such as being the site of the Stark or Zeeman effect, for example. If we would like to define the proper functions and proper values in space-time then we would find that the stationary states of that atom will be invariable and will be determined by the set of all the effects that it is subjected to during the entire course of time. That hardly seems acceptable.

In reality, even in the theory of relativity in its classical form today, the time and space variables are far from equivalent. The variable "time" always varies in the same sense in it, and the world-lines of all material objects are lines that have a positive sense, and that positive sense will always subtend an angle of at least 45° with the ct axis. In other words, space-time possesses an essential "polarization."

In the relativistic conception of things, an observer A considers the world-points that are contained in a certain three-dimensional section of space-time to be simultaneous and to correspond to the same value of its proper time. That is because that section cuts all of the world-lines that the observer A can cut out from its space of almost-independent units. However, such a cutting of the sense of world-lines would be impossible. There would be a sort of fibrous structure on space-time in the sense of time. It is that fibrous structure that we find to be inconvenient here, and we see that the difficulty has its roots in classical relativity itself.

4. A more restrictive form for the uncertainty relations (Bohr, Landau, and Peïerls). – If one introduces the relativistic idea that no action can propagate with a velocity that is greater than c then the relation $\Delta W \cdot \Delta t \geq h$ will lead us to state some new uncertainty relations. Those new relations, which are not contained in the general principles of the new mechanics in its non-relativistic form, must be added to those of Heisenberg and will augment the uncertainties that result from them.

Without entering into the details of the argument that leads to those new uncertainties, we shall meanwhile seek to show their origin. We put ourselves in the case in which no field is present.

In relativistic wave mechanics, the four parameters W, p_x, p_y, p_z that enter into the phase of a monochromatic plane wave are coupled by the relation of Einsteinian dynamics:

$$| p |^2 = \frac{W^2}{c^2} - m_0^2 c^2 , \tag{9}$$

from which, it will easily result that $\Delta | p | \geq \Delta W / c$.

However, we know that if one seeks to determine the state of an electron by means of an experiment that lasts for a duration of Δt then one can know that state with only an uncertainty of $h / \Delta t$; i.e., if the wave Ψ is decomposed into monochromatic waves then one will be concerned with a spectral interval of width:

$$\Delta \nu = \frac{\Delta W}{h} \geq \frac{1}{\Delta t} . \tag{10}$$

The values of $| p |$ that enter into the spectral decomposition of the wave Ψ will then occupy an interval:

$$\Delta | p | \geq \frac{\Delta W}{c} \geq \frac{h}{c \, \Delta t} . \tag{11}$$

One concludes from this that an observation or experiment of time duration Δt cannot lead to knowledge of the quantity of motion that has an uncertainty that is less than $h / c \, \Delta t$. That will yield an uncertainty relation with a new character.

One can also recover the inequality (11) in another manner. If an observer that is placed at a point in space would like to bound the length of a train of waves Ψ then he must make it pass through an opening that has been pierced through a screen and is closed by means of a shutter. He raises the shutter for a time interval of Δt and then puts it back in such a fashion as to bound the wave train that has crossed the screen during that time interval Δt. However, the front of that wave train cannot progress with a velocity that is greater than c, so the length of the transmitted wave train that crosses the screen will be equal to at most $c \, \Delta t$. Upon taking the z axis to be in the direction of propagation, the uncertainty in the z coordinate of the corpuscle of the transmitted wave will then be:

$$\Delta z \leq c \, \Delta t, \tag{12}$$

and from the Heisenberg relation:

$$\Delta p_z \geq \frac{h}{\Delta z} \geq \frac{h}{c \, \Delta t} , \tag{13}$$

and since one can confuse p and p_z here, one will recover the inequality (11).

In the first place, it might seem that the relation (11) is in opposition with the notion of a monochromatic plane wave because one has $\Delta p = 0$ for them. However, one can dismiss that objection by remarking that in order to rigorously have the right to associate a corpuscle with a monochromatic plane wave Ψ, one must be able to state that the corpuscle can be found at no particular point in space, and that statement can be justified only for an experiment of infinite time duration, since no means of investigation can explore space with a velocity that is greater than c. Meanwhile, it still remains from that objection that an observer can never represent his knowledge of the state of an electron by an unbounded monochromatic plane wave rigorously.

Certain arguments lead one to believe that for a body of proper mass m_0, the measurement of a length that is less than $h / m_0 c$ or of a time duration that is less than $h / m_0 c^2$ is illusory ([1]). That permits one to hope that if one can exclude the consideration of distances that are less than $h / m_0 c$ and time intervals that are less than $h / m_0 c^2$ from the theory of the electron as having no meaning then one might make the difficulty of negative energies disappear by means of that itself. Then again, one only hopes for that.

_____

([1]) Cf., SCHRÖDINGER, Annales de l'Institut H. Poincaré, *loc. cit.*