THE FUNDAMENTAL FIRST-ORDER EQUATIONS OF THE QUANTUM THEORY.

Submitted by.

J. W. WARREN,

For M. Sc. Degree. July 1947.
(Internal).

(1) INTRODUCTION.

The classical, non-relativistic, wave mechanics of L.de Broglie and Schroedinger has achieved considerable success in the interpretation of atomic phenomena. There are however several phenomena which remain unexplained by the elementary theory. The principle difficulties arise in the problems of fine structure in optical and X-Ray spectra, and in the Zeeman effect.

Using the methods of the old quantum theory Sommerfeld (ref.1) was able to deduce a formula giving the fine structure of the energy levels of hydrogen-like atoms. The electron was assumed to follow an elliptical path described by a total quantum number 'n' and an azimuthal quantum number 'k'. Account being taken of the variation of the mass of the electron with its velocity, as given by the theory of relativity, the energy of the orbit becomes

$$E_{nk} = -\frac{R N^2 h}{n^2}\left\{1 + \frac{a^2 N^2}{n^2}(n/k - \tfrac{3}{4})\right\} \quad \ldots\ldots \quad (1)$$

where R is Rydberg's constant and N is the atomic number of the element. The quantity 'a' is the fine structure constant given by

$$a = \frac{2\pi e^2}{hc} = \frac{1}{137} \text{ (e in O.E.S.U.)}$$

The formula (1) describes fairly accurately the fine structure of the spectra of atomic hydrogen, singly ionized helium etc. It is not exact, however, since it does not predict enough components. (ref. 2)

The energy levels given by non-relativistic wave mechanics are

$$E_{nL} = -\frac{R N^2 h}{n^2}\left\{1 + \frac{a^2 N^2}{n^2}\left(\frac{n}{L + \frac{1}{2}} - \frac{3}{4}\right)\right\} . \quad . \quad . \ (2)$$

where L = k -1.

The separation of the energy levels given by equation (2) is much larger than that given by equation (1) and is not in agreement with experiment. Hence wave mechanics in its original form fails to account satisfactorily for the fine structure of the spectra of hydrogen-like atoms. (ref. 3).

The multiplet structure of the spectral energy terms of atoms with more than one electron appeared to the early workers to be quite distinct from the fine structure shown by hydrogen-like atoms. Sommerfeld found it necessary to introduce an additional "inner" quantum number $j = L \pm \frac{1}{2}$ to account for the multiplet structure of both optical and X-Ray spectra. By this means he found a fine structure for X-Ray spectra and for the optical spectra of the alkali metals identical with that given by equation (1), except that the atomic number N is replaced by (N - z), where the term 'z' accounts for the screening of the nucleus by the inner orbital electrons.

Uhlenbeck and Goudsmit suggested a hypothesis to explain the signigifance of the new half-integral quantum number. They proposed to consider the electron as a spinning sphere of electricity possessing angular momentum $\frac{h}{4\pi}$ and a magnetic moment $B = \frac{he}{4\pi m_o c}$, equal to the Bohr magneton. The two values of j for each value of L correspond to opposite settings of the spin axis.

The ratio of the magnetic to the mechanical moment is $e/m_o c$, which is twice the corresponding ratio for an electronic orbit. This is the value that is necessary to explain the Zeeman effect and the Gyromagnetic Anomaly in ferromagnetic metals, (ref.4).

We thus find that the splitting of spectral lines in the case of hydrogen-like atoms came to be attributed to a relativity correction, whilst in the case of all other atoms it was attributed to the spin of the electron. The similarity of the results in the two cases thus appeared to be due to accident. (ref.5).

In this paper we shall first consider how the conception of electron spin is to be incorporated into wave mechanics. We shall then reformulate wave mechanical theory in accordance with the requirements of the special theory of relativity, and show that the spin and the magnetic moment of the electron arise quite naturally in the process. Finally we shall see that the exact relativistic theory of wave mechanics provides a satisfactory solution of the problems enumerated above.

(2) WAVE-MECHANICAL REPRESENTATION OF ANGULAR MOMENTUM.

The orbital moment of momentum or angular momentum of a particle about a point O, taken as origin of co-ordinates, is defined by the vector product

$$M = \left[r \times p \right] \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (3)$$

which has rectangular components

$$M_x = yp_z - zp_y$$
$$M_y = zp_x - xp_z$$
$$M_z = xp_y - yp_x$$

In wave mechanics these quantities will be replaced by operators (see paragraph 4) such that

$$\left.\begin{aligned} (M_x) &= \frac{h}{2\pi i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = \frac{h}{2\pi i}\frac{\partial}{\partial\phi_x} \\ (M_y) &= \frac{h}{2\pi i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) = \frac{h}{2\pi i}\frac{\partial}{\partial\phi_y} \\ (M_z) &= \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = \frac{h}{2\pi i}\frac{\partial}{\partial\phi_z} \end{aligned}\right\} \quad . \; . \; . \quad (4)$$

where the operators (M_ζ) are the wave mechanical equivalents of the quantities M_ζ, whilst the angles ϕ_ζ represent the azimuths about the corresponding axes.

It can be shown easily that the proper values of the operators (M_ζ) have the form $\frac{h}{2\pi}m$, where m is an integer. Thus $\frac{h}{2\pi}$ may be considered to be a fundamental unit of angular momentum.

Let us now consider the operator corresponding to the magnitude of M^2. We have

$$(M^2) = (M_x)^2 + (M_y)^2 + (M_z)^2$$

and the proper values of (M^2) can be shown to be $\frac{h^2}{4\pi^2} L(L + 1)$, where L is an integer.

The operator (M^2) commutes with the operators (M_j) but the latter do not commute with each other. Hence the simultaneous measurement of any two of the quantities M_j is subject to the restrictions imposed by Heisenberg's uncertainty principle. The relations of non-commutation between these operators are very important. To examine them we write the operators in terms of the unit $\frac{h}{2\pi}$,

$$(M_j) = \frac{h}{2\pi}\ m_j \qquad (j = x,y,z)$$

where the quantities m_j are operators which, from equation (4), are seen to obey the laws

$$\left.\begin{aligned} \left[m_x,\ m_y\right] &= -im_z \\ \left[m_y,\ m_z\right] &= -im_x \\ \left[m_z,\ m_x\right] &= -im_y \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (5)$$

It can be shown that the relationships (5) link the angular momenta with the group of spacial rotations, which obey an analogous law of non-commutation. (ref.6)

We must now consider the spin angular momentum. Let the spin of a particle have components S_x, S_y, S_z, represented by the operators (S_x), (S_y), (S_z).

The magnitude of S^2 will be represented by an operator such that

$$(S^2) = (S_x)^2 + (S_y)^2 + (S_z)^2$$

The spin operators are assumed to be related to the group of spacial rotations in the same way as are the orbital angular momentum operators. The theory of the spin operators may thus be deduced directly by analogy with our previous results. We go one step beyond the results for the orbital case in assuming the possibility of half-integral proper values. Thus we write the proper values of (S^2) as $\frac{h^2}{4\pi^2}s(s+1)$ where s may be integral or half-integral.

The operators (S_j) will be written

$$(S_j) = \frac{h}{2\pi} s_j \quad (j = x,y,z) \; . \; . \; . \; . \; . \; (6)$$

where the operators s_j obey the non-commutation rules

$$\left.\begin{aligned} [s_x, s_y] &= -is_z \\ [s_y, s_z] &= -is_x \\ [s_z, s_x] &= -is_y \end{aligned}\right\} \; . \; . \; . \; . \; . \; . \; . \; (7)$$

In the case of an electron the hypothesis of Uhlenbeck and Goudsmit assigns the values $\pm \frac{h}{4\pi}$ to each of the quantities S_j. We thus see that the proper values of s_j will in this case be $\pm \frac{1}{2}$.

We thus write

$$\sigma_j = 2\, s_j \qquad (j = x, y, z) \quad \ldots\ldots (8)$$

and seek for three operators σ_j which have the proper values ± 1 and are such that the operators s_j will satisfy the equations (7). This can be done most simply by assuming that the operators are matrices with two rows and two columns. As we shall see later in the analogous case in Dirac's theory the choice of matrices is largely arbitrary. We shall adopt the following set of three "Pauli matrices" which give a special prominance to the z-axis,

$$\sigma_x = \begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix}; \quad \sigma_y = \begin{Bmatrix} 0 & i \\ -i & 0 \end{Bmatrix}; \quad \sigma_z = \begin{Bmatrix} 1 & 0 \\ 0 & -1 \end{Bmatrix}; \quad \ldots\ldots (9)$$

These matrices which are clearly Hermitian can easily be shown to satisfy the conditions laid down. Their significance will become apparent in the next paragraph.

(3). PAULI'S THEORY.

We shall consider very briefly the attempt made by Pauli to incorporate the hypothesis of electron spin into wave-mechanics. (ref. 7).

Pauli's idea was to represent the two possibilities of the setting of an electron's spin in a given direction by the use of a wave function with two components ψ_1, and ψ_2.

We then write:

$$\left.\begin{aligned}\pi_+ &= \psi_1^*\psi_1 d\tau = |\psi_1|^2 d\tau \\ \pi_- &= \psi_2^*\psi_2 d\tau = |\psi_2|^2 d\tau\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (10)$$

where π_+ is the probability that the electron is in the element of volume $d\tau$ whilst the value of the spin angular momentum in the z-direction is $+\frac{h}{4\pi}$. Similarly π_- is the probability that the electron is in the element of volume $d\tau$ whilst the value of the spin in the z-direction is $-\frac{h}{4\pi}$.

The functions ψ_N are normalized by writing

$$\int_D \left\{|\psi_1|^2 + |\psi_2|^2\right\} d\tau = 1 \quad \ldots\ldots\ldots (11)$$

the integration being carried out over the whole domain denoted by D.

We have now to deduce the wave equation for the two-component wave function ψ, which we assume to have the form of a single column matrix,

$$\psi = \begin{Bmatrix}\psi_1 \\ \psi_2\end{Bmatrix}$$

The operation of a matrix A with two rows and two columns upon ψ is given by

$$A\psi_N = \sum_M A_{NM}\psi_M \quad (n,m = 1,2) \quad \ldots\ldots\ldots (12)$$

Now Schroedinger's wave equation can be written in the form

$$\left\{H + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right\}\psi = 0 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; (13)$$

where H is the wave mechanical Hamiltonian operator corresponding to the classical Hamiltonian function, We have, symbolically,

$$H = \frac{1}{2m}p^2 + U \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (14)$$

where p is the momentum operator and U represents the potential energy.

For an electron in a magnetic field K there will be an additional potential energy term U_M given by

$$U_M = -(K.\mu) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; (15)$$

where μ is the magnetic moment of the electron due to its spin. Using the value of μ given by the hypothesis of Uhlenbeck and Goudsmit, we obtain from (15), (6), (8),

$$U_M = -B\,(K.\sigma) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; (16)$$

where σ is the operator with components σ_x, σ_y, σ_z.

We now add U_M to a Hamiltonian generalised to operate upon the two-component function ψ, thus

$$\left\{\left(\frac{1}{2m}p^2 + U\right)\delta - B\,(K.\sigma)\right\}\psi = 0 \quad . \; . \; . \; . \; . \; (17)$$

where $\delta = \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix}$, is the unit matrix.

We thus obtain two equations for the components of ψ, which we write

$$\left.\begin{array}{l}\left\{\left(\frac{1}{2m}p^2 + U\right) - B\,K_z\right\}\psi_1 - \left\{K_x + iK_y\right\}\psi_2 = 0 \\ \\ \left\{\left(\frac{1}{2m}p^2 + U\right) + B\,K_z\right\}\psi_2 - \left\{K_x - iK_y\right\}\psi_1 = 0\end{array}\right\} \quad \ldots\ldots (18)$$

These equations take us a step beyond Schroedinger's elementary wave mechanics since account is now taken of the magnetism of the electron.

We shall not pursue the study of Pauli's theory any further. We do not expect it to be exact, since no account has been taken of the requirements of relativity, and in fact the theory does not give correct results for the fine structure of the spectra of hydrogen-like atoms.

The importance of Pauli's theory lies in the fact that it reveals the possibility of incorporating the hypothesis of spin and proper magnetic moment into wave mechanics, by using a wave function with more than one component. Dirac was guided by Pauli's work when he deduced his more exact theory, which will be developed in the succeeding paragraphs of this paper.

(4). THE RELATIVISTIC MOMENTUM OPERATORS.

We must now consider the relativistic generalization of wave mechanics.

The non-relativistic wave equation may be deduced by the analogy between the variation principles of Maupertius and Hamilton in mechanics, and Fermat's principle in geometrical optics.

The expression for the variation of the action of the path of a particle may be written in the form

$$\delta \int_1^2 \left\{ \sum_j p_j \, dq_j - W \, dt \right\} = 0 \quad (j = 1,2,3) \quad . \quad . \quad . \quad (19)$$

where W is the energy of the particle and p_j represents the component of momentum conjugate to the co-ordinate q_j.

Starting from equation (19) Schroedinger's wave equation can be deduced (ref.8) and may be written

$$\nabla^2 \psi + \frac{4\pi i m}{h}\frac{\partial \psi}{\partial t} - \frac{8\pi^2 m}{h^2} U \psi = 0 \quad . \quad . \quad . \quad . \quad . \quad (20)$$

or alternately, as in equation (13)

$$\left\{ H + \frac{h}{2\pi i}\frac{\partial}{\partial t} \right\} \psi = 0$$

We can write equation (14) for the Hamiltonian as

$$H = \left\{ \frac{1}{2m} \sum_j p_j^2 + U \right\} \quad (j = 1,2,3) \quad . \quad . \quad . \quad . \quad (21)$$

Equations (13) and (20) are seen to be equivalent if we replace the classical momentum components p_j in (21) by the operators given by

$$p_j \psi = \frac{h}{2\pi i}\frac{\partial \psi}{\partial q_j} \quad (j = 1,2,3) \quad . \quad . \quad . \quad . \quad (22)$$

In the special case where there is no potential field, equation (19) may be written in a form which is more in accordance with relativistic ideas by adopting Minkowski coordinates and momenta. Thus, putting $p_4 = imc$ and $q_4 = ict$ equation (19) may be written

$$\delta \int_1^2 \sum_\mu p_\mu dq_\mu = 0 \qquad (\mu = 1,2,3,4) \quad . \; . \; . \; (23)$$

From this expression we can deduce the wave equation for a free particle

$$\nabla^2 \psi + \frac{4\pi i m}{h} \frac{\partial \psi}{\partial t} = \frac{8\pi^2}{h^2} m_0^2 c^2 \psi \quad . \; . \; . \; . \; . \; . \; (24)$$

The energy in this case is $mc^2 = -icp_4$ and we can now extend equation (22) to all four components, thus

$$p_\mu \psi = \frac{h}{2\pi i} \frac{\partial \psi}{\partial q_\mu} \quad (\mu = 1,2,3,4) \quad . \; . \; . \; . \; . \; (25)$$

As can be seen from equations (13) and (21) the fourth of the relations (25) will not hold in a field of force. Thus in general we do not expect the equations (25) to be true in an electro-magnetic field. To determine the general form of these equations we consider the force on the particle, which is given in empty space by the well-known equation

$$f = e \left\{ J + \frac{1}{c} \left[v \times K \right] \right\} \quad . \; . \; . \; . \; . \; . \; . \; (26)$$

where e is the charge on the particle, J is the electric field and K is the magnetic field. These latter quantities depend on the scalar and vector potentials according to the equations

$$\left.\begin{aligned} J &= -\operatorname{grad} V - \frac{1}{c}\frac{\partial A}{\partial t} \\ K &= \operatorname{curl} A \end{aligned}\right\} \quad \ldots\ldots\ldots (27)$$

We substitute these expressions in (26) whilst introducing Minkowski co-ordinates and a four-vector potential given by

$$(A_1, A_2, A_3, A_4) \equiv (A_x, A_y, A_z, iV)$$

We thus find that the components of the "Newtonian" force* on the particle are given by

$$f_\mu = e/c \left\{ \frac{\partial A_\nu}{\partial q_\mu} - \frac{\partial A_\mu}{\partial q_\nu} \right\} \frac{dq_\nu}{dt} \quad \ldots\ldots\ldots (28)$$

Now a general expression for f_μ which is true in any system of co-ordinates is (ref.9)

$$f_\mu = \left\{ \frac{\partial p_\mu}{\partial q_\nu} - \frac{\partial p_\nu}{\partial q_\mu} \right\} \frac{dq_\nu}{dt} \quad \ldots\ldots\ldots (29)$$

Thus from (28) and (29),

$$\left\{ \frac{\partial P_\mu}{\partial q_\nu} - \frac{\partial P_\nu}{\partial q_\mu} \right\} \frac{dq_\nu}{dt} = 0 \quad \ldots\ldots\ldots (30)$$

where $$P_\mu = p_\mu + \frac{e}{c} A_\mu \quad \ldots\ldots\ldots (31)$$

From the analogy between (29) and (30) the variation

---

* $f_\mu = F_\mu \sqrt{1 - \frac{v^2}{c^2}}$; F_μ is a component of the Minkowski force.

principle may be written as

$$\delta \int_1^2 P_\mu \, dq_\mu = 0 \qquad (\mu = 1,2,3,4) \; . \; . \; . \; (32)$$

The relativistic wave equation is thus to be deduced from equation (32) instead of from equation (19). The relations analogous to (22) which hold in an electro-magnetic field are (ref. 10),

$$P_\mu \psi = \frac{h}{2\pi i} \frac{\partial \psi}{\partial q_\mu} \; (\mu = 1,2,3,4) \; . \; . \; . \; . \; . \; (33)$$

$$\left.\begin{array}{lll} \text{i.e.} & (p_j + \frac{e}{c} A_j)\psi = \frac{h}{2\pi i} \frac{\partial \psi}{\partial q_j} & (j = 1,2,3) \\ \text{and} & (p_4 + \frac{e}{c} A_4)\psi = \frac{h}{2\pi i} \frac{\partial \psi}{\partial ict} & \\ \text{or} & (mc + \frac{e}{c} V)\psi = \frac{-h}{2\pi i c} \frac{\partial \psi}{\partial t} & \end{array}\right\} \; . \; . \; . \; . \; (33)'$$

---

(5). A RELATIVISTIC SECOND ORDER WAVE EQUATION.

In the special theory of relativity we have the well - known relationship between the momenta:

$$m^2 c^2 - \sum_j p_j^2 - m_0^2 c^2 = 0 \; . \; . \; . \; . \; . \; . \; . \; (34)$$

If we regard the momenta as operators acting upon a wave function ψ we may write

$$\left\{ p_t^2 - \sum_j p_j^2 - p_0^2 \right\} \psi = 0 \qquad . \; . \; . \; . \; . \; . \; (35)$$

where $p_t = mc$, and $p_0 = m_0 c$.

Substituting the operators given by equations (33)' we obtain

$$\frac{1}{c^2}\left\{\frac{h}{2\pi i}\frac{\partial}{\partial t} + eV\right\}^2\psi - \sum_j\left\{\frac{h}{2\pi i}\frac{\partial}{\partial q_j} - \frac{eA_j}{c}\right\}^2\psi = p_0^2\psi \quad . \; . \; . \; (36)$$

We shall now deduce the form which must be taken by the "probability density" in accordance with the wave equation (36).

The equation which is the complex conjugate of (36) is

$$\frac{1}{c^2}\left\{-\frac{h}{2\pi i}\frac{\partial}{\partial t} + eV\right\}^2\psi^* - \sum_j\left\{-\frac{h}{2\pi i}\frac{\partial}{\partial q_j} - \frac{eA_j}{c}\right\}^2\psi^* = p_0^2\psi^* \quad . \; . \; . \; (36)^*$$

Equation (36) is now multiplied in front by ψ^* and equation $(36)^*$ by ψ. Then, subtracting,

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

$$+ \operatorname{div}\left\{\psi^* \operatorname{grad}\psi - \psi \operatorname{grad}\psi^* - \frac{4\pi i e}{h c} A\psi\psi^*\right\} = 0$$

or

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho v) = 0 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; (37)$$

$$\left.\begin{aligned} \text{where} \quad \rho &= \frac{h}{4\pi i m c}\left\{\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right\} - \frac{e}{m_0 c^2} V\psi\psi^* \\ \text{and} \quad \rho v &= \frac{h}{4\pi i m_0}\left\{\psi \operatorname{grad}\psi^* - \psi^* \operatorname{grad}\psi\right\} - \frac{e}{m_0 c} A\psi\psi^* \end{aligned}\right\} . \quad (38)$$

Equation (37) is the familiar equation which expresses the conservation of a quantity of density ρ moving with a velocity v.

Dirac has criticised the second order equation (36) and in particular the form of the probability density ρ derived from it. (ref. 11).

On considering the expression (38) for ρ we notice that the value is not necessarily positive, whilst a negative density is meaningless.

An even more serious difficulty is as follows. The general principles of wave mechanics demand (ref. 12) that the probability density should be given by $\rho = \psi^* \psi$ no matter what the form of the wave equation. (If there were several components ψ_N of ψ we would have $\rho = \sum_N \psi_N^* \psi_N$). If we were to adopt this value for ρ however, the equation (37) would not be satisfied. Hence the probability of locating a particle, and also the electric charge, would not be conserved.

These difficulties do not arise in the case of the approximate equation (20). Examination of the manner by which equation (37) was derived shows us immediately that the difficulty is due to the presence of the second order term $\frac{\partial^2 \psi}{\partial t^2}$. Hence we deduce that the true wave equation must be of the first order in $\frac{\partial}{\partial t}$. Now in relativity physics there is always symmetry between the space and time co-ordinates, hence we conclude that the wave equation should be of the first order in all the variables.

(6) THE FIRST ORDER WAVE EQUATIONS.

Consider the basic equation (35), which is true in zero field,

$$\left\{p_t^2 - \sum_j p_j^2 - p_0^2\right\}\psi = 0$$

This equation is to be considered as being derived from the unknown true wave equations which are of the first order.

Dirac assumed that the first order equations could be represented in the form

$$\left\{p_t + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_0 p_0\right\}\psi = 0 \qquad \ldots\ldots (39)$$

where the quantities α_j are Hermitian matrices with 'n' rows and 'n' columns. If 'n' is less than four it is impossible to find a set of 4 suitable matrices. Dirac assumes that the correct value of 'n' is 4, and that the matrices α_j operate upon a four-component function ψ such that

$$\alpha_j \psi_n = \sum_l (\alpha_j)_{nl} \psi_l \quad (n, l, = 1,2,3,4) \qquad \ldots (40)$$

thus we obtain four first order equations

$$\left\{p_t + \sum_1^3 \alpha_j p_j + \alpha_0 p_0\right\}\psi_n = 0 \qquad \ldots\ldots (41)$$

We now multiply equation (41) in front by the operator

$$\left\{p_t - \sum_j \alpha_j p_j - \alpha_0 p_0\right\}$$

This gives us the equation (35) for each of the components of ψ provided the matrices α_j and α_0 obey the anti-commutation law

$$\alpha_N \alpha_M + \alpha_M \alpha_N = 2\,\delta_{NM} \quad \dots\dots\dots\dots \quad (42)$$

where
$$\delta_{NM} = 1, \quad \text{if } m = n$$
$$= 0, \quad \text{if } m \neq n$$

The choice of matrices fulfilling the conditions (42) is largely arbitrary. It must be shown that this arbitrariness leads to no difficulties in the application of the theory.

We note that if we take two different sets of Hermitian matrices α_N and α_N' each satisfying equations (42) it is always possible to find a unitary matrix S with four rows and four columns such that

$$\alpha_N' = S^{-1}\alpha_N S = S^{+}\alpha_N\, S$$

Dirac's equations may be written in terms of the matrices α_N', the corresponding wave functions being ψ_N',

$$\left\{p_t + \alpha_1' p_1 + \alpha_2' p_2 + \alpha_3' p_3 + \alpha_0' p_0\right\} \psi_N' = 0$$

i.e. $$\left\{p_t + p_1 S^{-1}\alpha_1 S + p_2 S^{-1}\alpha_2 S + p_3 S^{-1}\alpha_3 S + p_0 S^{-1}\alpha_0 S\right\} \psi_N' = 0 \quad \dots \quad (43)$$

On multiplying in front by S we obtain Dirac's equation with the operators α_1, the wave function being now

$$\psi_N = S\,\psi_N'$$

Now all the quantities having any physical significance that are encountered in Dirac's theory have the form $\sum_N \psi_N^* A \psi_N$ (or the integrals of such expressions) where A is a linear and Hermitian operator containing the matrices α_N. On changing from a system in α_N to a system in α_N' we find that A becomes A' where

$$A' = S^{-1} A S$$

but

$$\sum_N \psi_N^* A \psi_N = \sum_N S^* \psi_N'^* A \; S \psi_N'$$

$$= \sum_N \psi_N'^* S^{-1} A \, S \psi_N' \quad (\text{since } S^{\dagger} = S^{-1})$$

Hence

$$\sum_N \psi_N^* A \, \psi_N = \sum_N \psi_N'^* A' \psi_N'$$

Thus the quantities having physical significance do not change their value on changing the matrices α_N. Therefore the arbitrary choice of these matrices does not affect the physical application of the theory.

The following set of matrices will be used throughout this paper:

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

$$\alpha_3 = \begin{Bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{Bmatrix}; \qquad \alpha_0 = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}$$

With this choice of matrices Dirac's equations can be written

$$\left.\begin{aligned} (p_t + p_0)\psi_1 + (p_1 + ip_2)\psi_4 + p_3\psi_3 = 0 \\ (p_t + p_0)\psi_2 + (p_1 - ip_2)\psi_3 - p_3\psi_4 = 0 \\ (p_t - p_0)\psi_3 + (p_1 + ip_2)\psi_2 + p_3\psi_1 = 0 \\ (p_t - p_0)\psi_4 + (p_1 - ip_2)\psi_1 - p_3\psi_2 = 0 \end{aligned}\right\} \quad \ldots\ldots (44)$$

There is a further degree of arbitrariness apparent in the choice of matrices. We have assumed tacitly that the quantity p_t should be multiplied by the unit matrix in equation (39). This assumption appears to give a special distinction to the time co-ordinate. We can transfer this pre-eminance to any other co-ordinate. For example, let us multiply equation (39) in front by α_0 . We thus obtain

$$\left\{\alpha_0 p_t + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 + p_0\right\}\psi = 0 \quad \ldots\ldots (45)$$

where $\beta_j = \alpha_0\alpha_j \qquad (j = 1,2,3)$

We see, however, that the matrices β_j are anti-Hermitian and therefore equations such as (45) seem less acceptable than the form (39).

This apparently special position of the time component is lost when we use Minkowski co-ordinates. The momentum p_t is replaced by $-ip_4$, and we can now write Dirac's equation (39) in a number of alternative forms like (45). In each case one of the matrices will be anti-Hermitian whilst the remainder are Hermitian or vice-versa.

(7) THE RELATIVISTIC INVARIANCE OF DIRAC'S EQUATIONS.

We shall employ Minkowski co-ordinates, the first order equations (41) being written

$$\left\{-ip_4 + \sum_j p_j\alpha_j + p_0\alpha_0\right\}\psi_N = 0 \quad \ldots\ldots\ldots (46)$$

Multiplying in front by $i\alpha_0$, we have

$$\left\{p_4\alpha_0 + \sum_j i\alpha_0\alpha_j p_j + ip_0\right\}\psi_N = 0$$

If we now introduce the matrices γ_N defined by

$$\gamma_1 = i\alpha_0\alpha_1 \; ; \quad \gamma_2 = i\alpha_0\alpha_2 \; ; \quad \gamma_3 = i\alpha_0\alpha_3 \; ; \quad \gamma_4 = \alpha_0 \; \ldots (47)$$

we find $\qquad \gamma_N\gamma_M + \gamma_M\gamma_N = 2\,\delta_{NM} \quad \ldots\ldots\ldots\ldots\ldots (48)$

where
$$\delta_{NM} = 1 \quad \text{if } n = m$$
$$= 0 \quad \text{if } n \neq m$$

The matrices γ_N are easily seen to be Hermitian.

We are now able to write Dirac's equations in the more symmetrical form

$$\left\{\sum_{\mu} p_\mu \gamma_\mu \quad + ip_0\right\} \psi_N = 0 \quad (\mu = 1,2,3,4) \quad . \; . \; . \; . \; (49)$$

We now subject the co-ordinates to a general Lorentz transformation, that is, a rotation in Minkowski space.

A four-vector such as the momentum transforms according to the law

$$p_M = \sum_N Q_{MN} p'_N \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (50)$$

Q_{MN} is the typical component of a matrix Q with four rows and four columns. Since we are concerned with rectangular co-ordinates we can write

$$\sum_L Q_{LM} \quad Q_{LN} = \delta_{MN} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (51)$$

The fourth component of a four-vector in the Minkowski continuum being imaginary, the matrix Q is not real in general.

After the change of axes the equation (49) becomes

$$\left\{\sum_{\mu} \gamma_\mu \sum_{\nu} Q_{\mu\nu} p'_\nu \; + \; ip_0\right\} \psi_N = 0 \quad . \; . \; . \; . \; . \; . \quad (52)$$

Now if we write

$$\gamma'_\nu = \sum_{\mu} Q_{\mu\nu} \gamma_\mu \qquad . \; . \; . \; . \; . \; . \quad (53)$$

equation (52) becomes

$$\left\{\sum_{\nu} \gamma'_\nu p'_\nu \; + \; ip_0\right\} \psi_N = 0 \qquad . \; . \; . \; . \; . \; . \quad (54)$$

Since the matrix Q is not purely real we see that the matrices γ_ν' are not in general Hermitian. They can, however, be shown to obey the same law of non-commutation as the matrices γ_μ .

After Neumann, we shall endeavour to represent the transformation (53) in the form

$$\gamma_\mu' = \Lambda^{-1}\gamma_\mu\Lambda \quad \ldots\ldots\ldots\ldots (55)$$

We notice that since γ_μ' is not Hermitian the operator Λ is not unitary, that is $\Lambda^{\dagger} \neq \Lambda^{-1}$.

It is necessary to show that the transformation (55) is possible. However we see that if

$$\gamma_\mu' = \Lambda_1^{-1}\gamma_\mu\Lambda_1 \quad \text{and} \quad \gamma_\mu'' = \Lambda_2^{-1}\gamma_\mu'\Lambda_2$$

then $$\gamma_\mu'' = \Lambda_2^{-1}\Lambda_1^{-1}\gamma_\mu\Lambda_1\Lambda_2 = (\Lambda_1\Lambda_2)^{-1}\gamma_\mu(\Lambda_1\Lambda_2)$$

Thus if the relation (55) is true for an infinitesimal rotation of the axes it is clearly true for a finite rotation.

Now in the case of an infinitesimal rotation we can put

$$Q_{\mu\nu} = \delta_{\mu\nu} + e_{\mu\nu} \quad \ldots\ldots\ldots (56)$$

where $e_{\mu\nu}$ is a very small quantity. To satisfy the condition of orthogonality (51) we have $e_{\mu\nu} = - e_{\nu\mu}$.

In the case of infinitesimal rotations the matrix Λ is very nearly equal to the unit matrix, and we assume

$$\Lambda = 1 + \sum_{\mu\nu} e_{\mu\nu} T^{\mu\nu}$$

hence $$\Lambda^{-1} = 1 - \sum_{\mu\nu} e_{\mu\nu} T^{\mu\nu} \quad \ldots\ldots (57)$$

where $T^{\mu\nu}$ is an anti-symmetric matrix which must be chosen so that

$$\sum_\mu Q_{\mu\nu}\gamma_\mu = \Lambda^{-1}\gamma_\nu\Lambda$$

Hence from (56) and (57),

$$\gamma_\nu + \sum_\mu e_{\mu\nu}\gamma_\mu = \left\{1 - \sum_{\alpha\beta} e_{\alpha\beta}T^{\alpha\beta}\right\}\gamma_\nu\left\{1 + \sum_{\alpha\beta} e_{\alpha\beta}T^{\alpha\beta}\right\}$$

$$= \gamma_\nu + \sum_{\alpha\beta} e_{\alpha\beta}\left\{\gamma_\nu T^{\alpha\beta} - T^{\alpha\beta}\gamma_\nu\right\}$$

Thus $$\sum_\mu e_{\mu\nu}\gamma_\mu = \sum_{\alpha\beta} e_{\alpha\beta}\left\{\gamma_\nu T^{\alpha\beta} - T^{\alpha\beta}\gamma_\nu\right\}$$

which gives us

$$T^{\alpha\beta} = -T^{\beta\alpha} = -\tfrac{1}{4}\gamma_\alpha\gamma_\beta$$

We are thus able to find matrices Λ for infinitesimal rotations. Hence it is possible to satisfy the relation (55) for all rotations.

Substituting from (55) in (54) we obtain

$$\sum_\nu \Lambda^{-1}\gamma_\nu\Lambda p_\nu' + ip_0\Big\}\psi = 0$$

multiplying in front by Λ we obtain Dirac's equations in the new co-ordinates,

$$\left\{\sum_\nu \gamma_\nu p_\nu' + ip_0\right\}\psi' = 0 \qquad \ldots\ldots (58)$$

where $$\psi' = \Lambda\psi$$

that is $$\psi_N' = \sum_M \Lambda_{NM}\psi_M. \quad \ldots\ldots\ldots \quad (59)$$

We thus see that Dirac's equations are invariant to a Lorentz transformation. The matrices γ_ν remain unchanged on changing the axes whilst the wave functions ψ_N transform according to the law (59). It must be noticed that ψ does not transform as a four-vector.

(8) THE FORMALISM OF DIRAC'S THEORY.

The formalism of Dirac's theory is analogous to that of non-relativistic wave mechanics. We must, however, always make a summation over the suffixes of ψ . There are also operators such as α_j , which do not appear in the older wave mechanics, which operate upon the suffixes of ψ .

At the very foundation of the new theory we have considered the form of the probability density ρ as derived from the second order equation (36), and noted Dirac's criticism of the result given by equation (38). We must now consider the corresponding expression in the new theory. Writing Dirac's equation and the complex conjugate equation we have

$$\left\{p_t + \sum_j \alpha_j p_j + \alpha_0 p_0\right\}\psi_N = 0$$

$$\left\{p_t^* + \sum_j \alpha_j^* p_j^* + \alpha_0^* p_0^*\right\}\psi_N^* = 0$$

Multiplying the first of these equations in front by ψ_N^* and the second by ψ_N and subtracting, we obtain, on making a summation over the suffixes of ψ,

$$\frac{\partial \rho}{\partial t} + \operatorname{div} j = 0 \quad \ldots\ldots\ldots \quad (60)$$

where

$$\left.\begin{aligned} \rho &= \sum_N \psi_N^* \psi_N \\ j &= -c \sum_N \psi_N^* \alpha \psi_N \end{aligned}\right\} \quad \ldots\ldots \quad (61)$$

where α is the operator with components $\alpha_1, \alpha_2, \alpha_3$.

Equation (60) represents the conservation of the probability of locating the particle. The flow vector $j = \rho v$, where v is the velocity of flow represented by $- c\alpha$.

The wave function ψ is normalised by writing

$$\left.\begin{aligned} \int_D \rho \, d\tau &= 1 \\ \text{i.e.} \quad \int_D \sum_N \psi_N^* \psi_N \, d\tau &= 1 \end{aligned}\right\} \quad \ldots\ldots \quad (62)$$

where $d\tau$ is an element of volume within the domain D.

With every quantity having physical significance will be associated a linear and Hermitian operator A, which will in general operate both on the co-ordinates and on the suffixes of ψ.

The observable values of the quantity associated with A will be the proper values 'a' of the equation

$$A \phi_N = a \phi_N \qquad (n = 1,2,3,4) \ \ldots \ (63)$$

There is thus a proper function ϕ_α with four components

$(\phi_\alpha)_N$ corresponding to the proper value a_α. In the non-degenerate case the proper functions will be orthogonal. If there is degeneracy we can make linear combinations of the degenerate proper functions such that the resulting functions are orthogonal. This in general we can write

$$\int_D \sum_N (\phi_\alpha)_N^* \, (\phi_\beta)_N \, d\tau = 0 \quad (\alpha \neq \beta) \quad \ldots\ldots \quad (64)$$

and the functions are normalized if we have

$$\int_D \sum_N (\phi_\alpha)_N^* (\phi_\alpha)_N \, d\tau = 1 \quad \ldots\ldots\ldots \quad (65)$$

Each of the components ψ_N of the wave function can be expanded as a series in terms of the proper functions $(\phi_\alpha)_N$, thus

$$\psi_N = \sum_\alpha c_\alpha (\phi_\alpha)_N \quad \ldots\ldots\ldots \quad (66)$$

where the terms c_α are constants, the squares of their moduli representing the probability of observing the corresponding proper values. Hence the mean value of A is given by

$$\bar{A} = \sum_\alpha a_\alpha |c_\alpha|^2 = \int_D \sum_N \psi_N^* A \psi_N \, d\tau \quad \ldots\ldots \quad (67)$$

Also the components of the matrix A are given by

$$A_{\alpha\beta} = \int_D \sum_N (\psi_\alpha)_N^* A \, (\psi_\beta)_N \, d\tau \quad \ldots\ldots \quad (68)$$

(9) THE CONSERVATION OF ANGULAR MOMENTUM.

We must now consider the meaning of the constancy or conservation of a quantity in our new wave mechanics. If we differentiate with respect to the time the matrix element A as expressed by equation (68) we obtain

$$\frac{dA_{\alpha\beta}}{dt} = \int_D \sum_N (\psi_\alpha)_N^* \left\{ \frac{\partial A}{\partial t} - \frac{2\pi i}{h} (A\,H - H\,A) \right\} (\psi_\beta)_N \, d\tau \quad . \, . (69)$$

$\frac{\partial A}{\partial t}$ is the partial differential coefficient of A obtained by formal derivation with respect to t, and H is the Hamiltonian.

If the quantities $A_{\alpha\beta}$ are to be conserved $\frac{dA_{\alpha\beta}}{dt}$ must be zero and hence the integral in (69) must vanish for all values of ψ, thus

$$\frac{\partial A}{\partial t} - \frac{2\pi i}{h} (A\,H - H\,A) = 0 \quad . \; . \; . \; . \; . \quad (70)$$

now, symbolically,

$$\frac{\partial A}{\partial t} = \left(\frac{\partial}{\partial t} A - A \frac{\partial}{\partial t} \right)$$

hence (70) becomes

$$R\,A - A\,R = \left[R\,,\,A \right] = 0 \quad . \; . \; . \; . \; . \quad (71)$$

where $$R = \left\{ H + \frac{h}{2\pi i} \frac{\partial}{\partial t} \right\} \quad . \; . \; . \; . \; . \; . \quad (72)$$

Hence a quantity will be conserved if the corresponding operator A commutes with R, or , in the case where A does not depend explicitly upon the time, if it commutes with H.

Applying this test to the energy and momentum operators we can demonstrate the conservation of these quantities. The moment of momentum, however, provides a special problem.

The angular momentum of a particle about the z-axis, M_z, is represented by the operator

$$(M_z) = xp_y - yp_x = \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \quad . \ . \ . \ (73)$$

Consider the case where the field of force is cylindrically symmetrical about the z-axis. In this case it is found that M_z is conserved according to the old wave mechanics where the Hamiltonian is of the second order with respect to the space co-ordinates. In the new theory this no longer applies.

Let us consider the value of $\left[R\ ,\ (M_z)\right]$, (M_z) must commute with $\frac{h}{2\pi i}\frac{\partial}{\partial t}$, $\alpha_3\frac{h}{2\pi i}\frac{\partial}{\partial x}$, $\alpha_0\, p_0$ and with $\frac{e}{c}V$ since there is cylindrical symmetry.

Now the other terms give

$$\left[R\ ,\ (M_z)\right] = \left[\left(\alpha_1\frac{h}{2\pi i}\frac{\partial}{\partial x} + \alpha_2\frac{h}{2\pi i}\frac{\partial}{\partial y}\right),\ \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\right]$$

hence $$\left[R\ ,\ (M_z)\right] = \frac{h^2}{4\pi^2}\left\{\alpha_2\frac{\partial}{\partial x} - \alpha_1\frac{\partial}{\partial y}\right\} \quad . \ . \ . \ . \ . \ . \ (74)$$

Hence (M_z) does not commute with R. It is however possible to find an operator containing (M_z) which does commute with R.

Consider the operator

$$(N_z) = (M_z) + i\,\alpha_1\alpha_2\frac{h}{4\pi}$$

on evaluating the commutator $\left[R\ ,\ (N_z)\right]$ we find

$$\left[R\ ,\ (N_z)\right] = R\,(N_z) - (N_z)R = 0 \quad \ldots\ldots\ldots \quad (75)$$

Hence the quantity N_z corresponding to the operator (N_z) is conserved. The moments of momentum about the other axes may be treated similarly and we find finally that a quantity N is conserved where

$$N = M + S \quad \ldots\ldots\ldots \quad (76)$$

S is the spin angular momentum and is represented by

$$(S) = \zeta \frac{h}{4\pi} \quad \ldots\ldots\ldots \quad (77)$$

where ζ is a Hermitian operator with components

$$\zeta_x = i\alpha_2\alpha_3 ; \quad \zeta_y = i\alpha_3\alpha_1 \ ; \quad \zeta_z - i\alpha_1\alpha_2 . \quad \ldots \quad (78)$$

It can be shown easily that the components of ζ have proper values ± 1. We thus see that the total angular momentum is conserved if we add to the orbital moment of momentum of a particle the spin moment of momentum, whose value measured in any direction is $\pm\frac{h}{4\pi}$. As we saw in paragraphs (1) and (2), this is the value that was proposed for the spin of the electron by Uhlenbeck and Goudsmit, in their semi-empirical attempt to solve the problems of fine-structure and the Zeeman effect. The remarkable feature of Dirac's theory is the fact that the correct value of the spin of the electron is deduced automatically in the derivation of a theory of relativistic wave mechanics. We note, however, the existence of particles (Photons, Mesons (ref.13)) having other values for their spin momenta. Here is apparently a limitation to the application of Dirac's equations in their present form.

Let us now calculate the mean value $\bar{N}_z$ of N_z . This is given by equation (67) in the form

$$\bar{N}_z = \int_D \sum_N \overset{*}{\psi}_N (N_z) \psi_N \, d\tau$$

$$\therefore \bar{N}_z = \int_D \sum_N \overset{*}{\psi}_N \left\{ (M_z) + \frac{h}{4\pi} i\alpha_1\alpha_2 \right\} \psi_N \, d\tau$$

$$= \bar{M}_z + \frac{h}{4\pi} \int_D \left\{ \overset{*}{\psi}_1 \psi_1 - \overset{*}{\psi}_2 \psi_2 + \overset{*}{\psi}_3 \psi_3 - \overset{*}{\psi}_4 \psi_4 \right\} d\tau$$

i.e. $$\bar{N}_z = \bar{M}_z + \frac{h}{4\pi} \int_D \left\{ |\psi_1|^2 + |\psi_3|^2 \right\} d\tau - \frac{h}{4\pi} \int_D \left\{ |\psi_2|^2 + |\psi_4|^2 \right\} d\tau \quad . . . (79)$$

We interpret this equation by saying that the probability that the value of the spin in the z-direction is $+\frac{h}{4\pi}$ is given by

$$\Pi_+ = \int_D \left\{ |\psi_1|^2 + |\psi_3|^2 \right\} d\tau$$

whilst the probability that the value is $-\frac{h}{4\pi}$ is given by

$$\Pi_- = \int_D \left\{ |\psi_2|^2 + |\psi_4|^2 \right\} d\tau$$

(10) DIRAC'S EQUATIONS IN AN ELECTRO-MAGNETIC FIELD.

We shall now derive the second order equation for a particle in an electro-magnetic field, on the basis of Dirac's first order equations.

We operate with

$$\left\{p_t - \sum_j \alpha_j p_j - \alpha_o p_o\right\}$$

upon the equation (39) which we write

$$\left\{p_t + \sum_j \alpha_j p_j + \alpha_o p_o\right\}\psi = 0$$

Taking account of the non-commutation relations (42) we obtain

$$\left\{p_t^2 - \sum_j p_j^2 - p_o^2 + \sum_j \alpha_j (p_t p_j - p_j p_t) - \sum_{j \neq k} (\alpha_j \alpha_k p_j p_k - \alpha_k \alpha_j p_k p_j)\right\}\psi = 0 \quad \ldots \quad (80)$$

The two summations can be expressed in terms of the electric and magnetic field strengths by using equations (33)' and (27), then

$$\left.\begin{aligned} \sum_j \alpha_j (p_t p_j - p_j p_t) &= -\frac{he}{2\pi i c}\left\{\alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3\right\} \\ \sum_{j \neq k} (\alpha_j \alpha_k p_j p_k - \alpha_k \alpha_j p_k p_j) &= -\frac{he}{2\pi i c}\left\{\alpha_2 \alpha_3 K_1 + \alpha_3 \alpha_1 K_2 + \alpha_1 \alpha_2 K_3\right\} \end{aligned}\right\} \quad \ldots \quad (81)$$

Substituting these values in (80) the second order equation takes the form

$$\left\{p_t^2 - \sum_j p_j^2 - p_o^2 - \frac{he}{2\pi i c}\left[(\alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3) - (\alpha_2 \alpha_3 K_1 + \alpha_3 \alpha_1 K_2 + \alpha_1 \alpha_2 K_3)\right]\right\}\psi = 0 \quad \ldots \quad (82)$$

The novel feature of this equation lies in the terms in square brackets. Their interpretation becomes apparent if we write Schroedinger's equation (20) as follows

$$\left\{2\ m\ c\ p_t - \sum_j p_j^2 - 2\ m\ U\right\}\psi = 0 \qquad \ldots\ldots (83)$$

On comparing the last two equations we see that the terms in square brackets in (82) correspond to a potential energy term U multiplied by 2 m. Assuming that the mass appropriate to the new theory is the proper mass $m_\circ$, we obtain the two potential energy terms.

$$U_M = \frac{he}{4\pi m_\circ c}\, i\left\{\alpha_2\alpha_3 K_1 + \alpha_3\alpha_1 K_2 + \alpha_1\alpha_2 K_3\right\} \qquad \ldots\ldots (84)$$

$$U_E = \frac{he}{4\pi m_\circ c}\, i\left\{\alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3\right\} \qquad \ldots\ldots (85)$$

These terms may be considered to be derived from a proper magnetic moment X and a proper electric moment Y of the particle, whose components are given by

$$\left.\begin{aligned} X_1 &= B\ i\alpha_2\alpha_3\ ; \quad X_2 = B\ i\alpha_3\alpha_1\ ; \quad X_3 = B\ i\alpha_1\alpha_2\ . \\ Y_1 &= B\ i\alpha_1\ ; \quad Y_2 = B\ i\alpha_2\ ; \quad Y_3 = B\ i\alpha_3\ . \end{aligned}\right\} \ldots\ldots (86)$$

Here B stands for the quantity $\frac{he}{4\pi m_\circ c}$, which in the cases of the electron and positron is equal to the Bohr magneton. The first set of equations (86) may be written

$$X = B\,\xi = \frac{e}{m_\circ c}(S) \qquad \ldots\ldots\ldots (87)$$

where (S) is the spin angular momentum defined by equation (77) and ξ is the Hermitian operator defined by equations (78). The ratio of the proper magnetic moment to the spin, given by equation (87), agrees with the value proposed for the electron by Uhlenbeck and Goudsmit.

Dirac's theory thus gives an account of the spin and proper magnetism of the electron. It appears that the results obtained above should apply to all elementary particles, provided that we insert the appropriate values of e and m_o. We must note that experimental results obtained for protons and neutrons do not agree with equation (87) (ref. 14).

The equations (86) indicate that in addition to the real magnetic moment of a particle there is an imaginary electric moment. Frenkel (ref.15) has shown that in the case of a moving particle the real magnetic moment will give rise to an additional real electric moment, whilst the imaginary electric moment gives rise to an imaginary magnetic moment.

The derivation of the electric and magnetic moments given above is not entirely satisfactory. In comparing equations (82) and (83) we assumed that the quantity m_o is the true representative of mass in the new theory. Now we have seen in equation (61) that the "classical" velocity 'v' is replaced in Dirac's theory by a very different quantity, that is the operator $-c\alpha$. We might well suppose that a similar difficulty would arise in the representation of mass. That this is indeed the case is indicated by a calculation due to Gordon.(ref.16)

His results, which we shall not consider in detail, indicate that the magnetic moment should be written

$$X = B\alpha_0 \xi \quad \ldots\ldots\ldots\ldots \quad (88)$$

This result may be interpreted, according to Frenkel, (ref.17), by considering the similarity between Dirac's equation (39) and an analogous first order equation in special relativity mechanics. The latter may be deduced from the well-known expression

$$m\,c^2 = m_0 c^2 (1 - \tfrac{v^2}{c^2})^{-\frac{1}{2}} \quad \ldots\ldots\ldots \quad (89)$$

hence
$$m\,c^2 = m_0 c^2 (1 - \tfrac{v^2}{c^2})^{\frac{1}{2}} + m_0 v^2 (1 - \tfrac{v^2}{c^2})^{-\frac{1}{2}}$$

or
$$m\,c^2 = m_0 c^2 (1 - \tfrac{v^2}{c^2})^{\frac{1}{2}} + (p.v)$$

where p represents the momentum m v.

We thus obtain an expression for the energy

$$W = U + (p.v) + m_0 c^2 (1 - \tfrac{v^2}{c^2})^{\frac{1}{2}} \quad \ldots\ldots \quad (90)$$

Now Dirac's equation can be written in the standard form

$$(H + \frac{h}{2\pi i}\frac{\partial}{\partial t})\psi = 0$$

where the Hamiltonian operator H corresponds to the energy W, and may be written, from (33)' and (39),

$$H = U - (c\,p.\alpha) - m_0 c^2 \alpha_0 \quad \ldots\ldots\ldots\ldots \quad (91)$$

Hence, comparing (90) and (91) we find, as before, that in the new theory the velocity v is replaced by $-c\alpha$. We also find that the operator α_0 replaces the quantity $-(1-\frac{v^2}{c^2})^{\frac{1}{2}}$ that is, the ratio $-m_0/m$.

Frenkel then assumes that the mass in equation (83) is the relative mass m, which in the new theory will be replaced by $-m_0/\alpha_0$.

Thus on comparing equations (82) and (83) we find that the magnetic moment is given by (88) instead of by (87). More important still, the components of the electric moment become

$$Y_1 = B\,i\alpha_0\alpha_1; \qquad Y_2 = B\,i\alpha_0\alpha_2; \qquad Y_3 = B\,i\alpha_0\alpha_3 \;.\;.\;. \quad (92)$$

The significant feature of this result is the fact that, unlike equations (86), the equations (92) predict a real electric moment.

L.de Broglie (ref.18) obtains the results (88) and (92) by a somewhat different argument. He observes that in Dirac's equations (44) the sign of the quantity $p_0 (= m_0 c)$ is different in the first pair of equations compared with the second pair. If we operate on each of the four functions ψ_N with $-\alpha_0 m_0$ we obtain

$$\left.\begin{array}{ll} -\alpha_0 m_0\psi_1 = -m_0\psi_1\;; & -\alpha_0 m_0\psi_2 = -m_0\psi_2. \\ -\alpha_0 m_0\psi_3 = m_0\psi_3\;; & -\alpha_0 m_0\psi_4 = m_0\psi_4\;. \end{array}\right\} \;.\;.\;. \quad (93)$$

Thus a mass m_0 is used in conjunction with the functions ψ_3 and ψ_4, whilst a mass $-m_0$ is used in conjunction with the functions ψ_1 and ψ_2. The choice of sign is made so that the positive mass is used with the functions which, as we shall see in the next paragraph, predominate in the "classical" case of low velocities.

In comparing the equations (82) and (83) de Broglie assumes that we are concerned with the proper mass m_0, which he replaces by $-m_0 \alpha_0 = -m_0/\alpha_0$. Hence we again obtain the results (88) and (92).

These different arguments due to de Broglie and Frenkel are seen to lead to the same results. Their difference lies in the manner in which a result is deduced in the new wave mechanics by analogy with the known results in non-relativistic wave mechanics.

If we accept the arguments outlined in this paragraph we are led to the conclusion that the electron and other charged particles possess a real proper electric moment defined by equations (92).

We must note that Cattermole and Wilson (ref.19) have published a theory of the electron, based on Kaluza's theory of relativity, which gives no electric moment, real or imaginary.

(11). PLANE WAVES IN DIRAC'S THEORY.

Let us consider the equations (44) in the absence of any field of force. In this case the relations (33)' take the simple form

$$p_j \psi = \frac{h}{2\pi i}\frac{\partial \psi}{\partial q_j} \quad ; \qquad p_t \psi = -\frac{h}{2\pi i c}\frac{\partial \psi}{\partial t} \quad . \; . \; . \; . \quad (94)$$

Equations (44) then become

$$\left.\begin{aligned}
&\left\{-\frac{h}{2\pi i c}\frac{\partial}{\partial t} + m_o c\right\}\psi_1 + \left\{\frac{h}{2\pi i}\left(\frac{\partial}{\partial q_1} + i\frac{\partial}{\partial q_2}\right)\right\}\psi_4 + \frac{h}{2\pi i}\frac{\partial \psi_3}{\partial q_3} = 0\\
&\left\{-\frac{h}{2\pi i c}\frac{\partial}{\partial t} + m_o c\right\}\psi_2 + \left\{\frac{h}{2\pi i}\left(\frac{\partial}{\partial q_1} - i\frac{\partial}{\partial q_2}\right)\right\}\psi_3 - \frac{h}{2\pi i}\frac{\partial \psi_4}{\partial q_3} = 0\\
&\left\{-\frac{h}{2\pi i c}\frac{\partial}{\partial t} - m_o c\right\}\psi_3 + \left\{\frac{h}{2\pi i}\left(\frac{\partial}{\partial q_1} + i\frac{\partial}{\partial q_2}\right)\right\}\psi_2 + \frac{h}{2\pi i}\frac{\partial \psi_1}{\partial q_3} = 0\\
&\left\{-\frac{h}{2\pi i c}\frac{\partial}{\partial t} - m_o c\right\}\psi_4 + \left\{\frac{h}{2\pi i}\left(\frac{\partial}{\partial q_1} - i\frac{\partial}{\partial q_2}\right)\right\}\psi_1 - \frac{h}{2\pi i}\frac{\partial \psi_2}{\partial q_3} = 0
\end{aligned}\right\} \quad (95)$$

These equations have plane-wave solutions of the form

$$\psi_N = a_N \exp. -\frac{2\pi i}{h}\left\{Wt - (p.q)\right\} \quad . \; . \; . \; . \quad (96)$$

Substituting in equations (95),

$$\left.\begin{aligned}
&\left\{\frac{W}{c} + m_o c\right\} a_1 + \left\{p_1 + ip_2\right\} a_4 + p_3 a_3 = 0\\
&\left\{\frac{W}{c} + m_o c\right\} a_2 + \left\{p_1 - ip_2\right\} a_3 - p_3 a_4 = 0\\
&\left\{\frac{W}{c} - m_o c\right\} a_3 + \left\{p_1 + ip_2\right\} a_2 + p_3 a_1 = 0\\
&\left\{\frac{W}{c} - m\, c\right\} a_4 + \left\{p_1 - ip_2\right\} a_1 - p_3 a_2 = 0
\end{aligned}\right\} \quad . \; . \; . \quad (97)$$

The condition that these equations should be satisfied for non-zero values of a_N is that the determinant formed from the coefficients of the amplitudes a_N is zero. On evaluating this determinant we find

$$\frac{W^2}{c^2} = m_0^2 c^2 + p^2 \quad \ldots\ldots (98)$$

that is

$$W = \pm c\sqrt{m_0^2 c^2 + p^2} \quad \ldots\ldots (99)$$

The condition (98) is automatically satisfied, since it is the well-known relationship of special relativity mechanics.

We shall confine our attention for the moment to the positive solution in (99). Giving the arbitrary values A and B* to a_3 and a_4 respectively, we can determine the other two amplitudes, since

$$\left.\begin{aligned} a_1 &= \left\{\frac{p_3 A - (p_1 + ip_2)B}{\frac{W}{c} + m_0 c}\right\} \\ a_2 &= -\left\{\frac{p_3 B + (p_1 - ip_2)A}{\frac{W}{c} + m_0 c}\right\} \end{aligned}\right\} \quad \ldots\ldots (100)$$

In the case where the momenta p_j are small compared with $m_0 c$ we see that a_1 and a_2 are very small compared with a_3 and a_4. In a system of co-ordinates in which the particle is at rest the amplitudes a_1 and a_2 , and hence the wave functions ψ_1 and ψ_2, are zero. Hence in this limiting case we have a wave function with only two components.

---

* The solution B is not to be confused with the Bohr magneton.

(12) THE NEGATIVE ENERGY STATES.

The equation (98) can be satisfied by taking the negative sign for the square root in (99) instead of the positive. There are thus two possible values of the energy of a particle with a given momentum, a positive value which we shall call W_P, and a negative value which we shall call W_N, such that $W_N = - W_P$.

Taking the negative energy solutions we can obtain a set of values for the amplitudes a_N similar to (100) Thus

$$\left.\begin{aligned} a_1 &= C \;; \qquad a_2 = D \;; \\ a_3 &= \left\{\frac{p_3 C - (p_1 + ip_2)D}{\frac{W_N}{c} - m_0 c}\right\} \\ a_4 &= -\left\{\frac{p_3 D + (p_1 - ip_2)C}{\frac{W_N}{c} - m_0 c}\right\} \end{aligned}\right\} \quad \ldots\ldots \quad (101)$$

Clearly the wave functions ψ_3 and ψ_4 are negligible for the states of negative energy in the ordinary case where the momenta p_j are small compared with $m_0 c$.

The states of negative energy of a particle are an outstanding feature of Dirac's theory. The question does not arise in ordinary relativity mechanics, although the ambiguity of sign in equation (99) is of course found there also. There is however a discontinuity in the range of energy values of at least $2\, m_0 c^2$ between the positive energy states and the negative energy states. Therefore in ordinary physics a particle in a positive energy state could not pass into a state of negative energy. In the quantum theory we cannot eliminate the negative energy states in this way, since in this case a discontinuous change of energy is quite possible.

We must now see how such states of negative energy might be expected to arise in practice.

A particular case has been examined in detail by Klein (ref.20). He considered the incidence of a free electron upon a potential barrier, the change of potential energy on crossing the barrier being U. There are three cases:

(1) $U < mc^2 - m_0c^2$; there are both reflected and refracted waves, that is, there is a finite probability of the electron passing through the barrier.

(2) $mc^2 - m_0c^2 < U < mc^2 + m_0c^2$; the transmitted wave is imaginary.

(3) $U > mc^2 + m_0c^2$; there is a transmitted wave, which indicates a finite probability that the electron will pass through the barrier, in which case it will enter a state of negative energy.

Although it is impossible in practice to produce a static field sufficient to create the conditions of the third case, the theoretical probability is rather disturbing. We may note, however, that this probability becomes zero if we assume that the potential energy cannot change by an amount $m_0 c^2$ in a distance less than $h/m_0 c$.

If we endeavour to represent any general solution of the wave equation by the superposition of a series of "monochromatic" plane waves it is found that the waves corresponding to positive energy states alone do not form a "complete system" in the sense of Fourier's theorem. Such a complete system is only normally obtainable if the waves representing negative energy states are also included.

It can be shown that the system of waves representing positive energy states will form a complete system provided the dimensions of the wave train are considerably greater than $h/m_0 c$.

A detailed examination of the scattering of electromagnetic radiation by free electrons shows that the Dirac electron can only scatter such radiation provided it can enter states of negative energy. (ref.21). Thus it appears that the states of negative energy play an essential part in the theory, although no particle has ever been observed in such a state.

Finally we must notice that Dirac has made an attempt to incorporate the negative energy states into the theory in a very novel way. He assumes that these states are all normally occupied by negative electrons which produce no field and are in no way directly observable. By Pauli's exclusion principle we see that the difficulty proposed by Klein does not arise since all possible states are occupied, hence the electron cannot enter a negative energy state, in this case.

If a sufficient quantity of energy is given to an electron with negative energy it will be raised into a positive energy state and hence become observable. The "hole" in the system of negative energy states will display all the properties to be expected of a positive electron or positron, with positive energy. Before the discovery of the positron the "hole" was identified with a proton, but the difference in mass created a serious objection to this hypothesis.

If the "hole" theory is true we should expect pairs consisting of a positive and a negative electron to be formed, for example by the absorption of hard γ-rays of energy $> 2m_o c^2$ ($\lambda < \frac{1}{2}\frac{h}{m_o c}$). This phenomenon has been observed in cloud-chambers used in investigating cosmic rays. (ref.22).

Similarly an electron might be expected to fall into the "hole" represented by a positron, with the emission of radiation. This phenomenon also has been observed.

The formalism of the "hole" theory of positrons leads to several unresolved difficulties, and must be considered as provisional only.

The states of negative energy constitute a fundamental difficulty in relativistic wave mechanics. The solution appears to be associated with the existence of a minimum distance h/m_0c and a minimum proper time h/m_0c^2, associated with a particle of proper mass m_0.

(13) CONCLUSION.

C.G.Darwin and W.Gordon have succeeded in deducing the fine structure formula for the spectra of hydrogen-like atoms on the basis of Dirac's equations. (ref.23). Their formula gives for the energy levels, approximately

$$E_{Nj} = -\frac{R\,h\,N^2}{n^2}\left\{1 + \frac{a^2N^2}{n^2}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right)\right\} \quad \ldots\ldots (102)$$

where $j = L \pm \frac{1}{2}$,

The selection rules can be shown to be

$$\delta L = \pm 1\ ;\ \delta m = \pm 1, 0\ ;\ \delta j = \pm 1, 0. \quad \ldots (103)$$

where the quantum numbers L, m, j, have their usual significance.

The application of the formula (102) and the selection rules (103) gives results in complete agreement with experiment.

Similarly the application of Dirac's theory leads to correct results for the fine structure of X-Ray spectra, and for the Zeeman effect in hydrogen-like and alkali atoms.(ref.24)

Considerable difficulty is encountered in attempting to apply the first order equations to the many-body problem. We note, however, that this problem leads to great difficulty even in ordinary relativity mechanics.

There are other methods of developing and interpreting the first order equations than that discussed in this paper.

Darwin and Frenkel (ref.25) have drawn attention to the analogy between the first order equations of the quantum theory and Maxwell's electro-magnetic equations. In the electro-magnetic theory the first order equations were known first and the second order equation deduced from them. In the quantum theory we have the second order equation (35), which is true in zero field, and we have to find the corresponding first order equations. This treatment leads to the same results as Dirac's.

In more recent researches workers have tended to base their theories upon Kaluza's 5-dimensional relativity. The starting point of this theory lies in equations (30) and (32).(ref.26) It will be noticed that equation (30) is of the form of the equation of a geodesic, but does not represent a geodesic in the 4-dimensional continuum. By adopting a five dimensional co-ordinate system it is possible to make (30) represent a geodesic. The quantities P_μ are now components of a covariant vector in this continuum, the fifth component P_5 being identified with the charge on the particle multiplied by a suitable dimensional constant. The path of an electron is assumed to be a null-geodesic in the Kaluza space.

H.T.Flint has developed a theory of the electron on the basis of Kaluza's theory of relativity. (ref.27). He assumes that in the description of micro-physics the ordinary length and vectors of macro-physics are to be replaced by matrices. The wave function ψ is introduced as a "guaging function" for the continuum, on considering the change of the components of a vector in a parallel displacement. The detailed calculation leads to Dirac's equations of which it provides a new interpretation.

SYMBOLS.

Most of the symbols used are defined in the text or are of well-known conventional significance. To avoid any possible confusion some of the more important are defined below:-

$(X.Y)$ = the scalar product of X and Y.

$[X \times Y]$ = the vector product of X and Y.

$[X, Y]$ = $XY - YX$ = the commutator of X and Y.

e = the protonic charge = - the electronic charge.

m = the relative mass. In paragraph (2), m = an integer; in paragraph (13), m = a quantum number.

m_o = the proper mass of a particle.

p_N = a component of the momentum, or the corresponding wave mechanical operator.

P_N = $p_N + \frac{e}{c} A_N$ a component of the "extended momentum".

δ = a unit matrix. In paragraph (4), δ = "the variation of-" in paragraph (13), δ = "a change in the value of-"

REFERENCES.

(1) GENERAL.

This paper is based mainly upon the following works:

L. de Broglie: L'ÉLECTRON MAGNÉTIQUE: Hermann 1934.

L. de Broglie: Théorie Générale des PARTICULES A SPIN: Gauthier-Villars, 1943.

J. Frenkel: WAVE MECHANICS, part 2: Oxford, 1934.

W. Wilson: THEORETICAL PHYSICS, vol. 3: Methuen, 1940.

(2) REFERENCES IN TEXT.

(ref.1). L'ÉLECTRON MAGNÉTIQUE, p.10.

(2) Johnson: SPECTRA (Methuen 1941) p.36.
L'ÉLECTRON MAGNÉTIQUE, p.38.

(3) " " p.105.
Sommerfeld: WAVE MECHANICS, (Methuen 1930), p.114.

(4) THEORETICAL PHYSICS: chap.XIV.
L'ÉLECTRON MAGNÉTIQUE, chap. IV.
H.Wilson: MODERN PHYSICS, (Blackie 1937), p.184.

(5) Sommerfeld: WAVE MECHANICS, p.118.

(6) PARTICULES A SPIN, chap. 111.

(ref. 7). L'ÉLECTRON MAGNÉTIQUE, chap.IX.
Frenkel: WAVE MECHANICS, p.279.

(8) H.T.Flint: WAVE MECHANICS, (Methuen 1943).

(9) THEORETICAL PHYSICS, p.61 and p.68-69.

(10) " " p.67-69.
Weyl: THE THEORY OF GROUPS AND QUANTUM MECHANICS (Methuen, 1931) p.98-100.
For a more elementary treatment see McCREA: RELATIVITY PHYSICS, (Methuen 1935) p.57.

(11) L'ÉLECTRON MAGNÉTIQUE, p.133-135.

(12) " " p.67.

(13) PARTICULES A SPIN.

(14) L. de Broglie: DE LA MÉCANIQUE ONDULATOIRE A LA THÉORIE DU NOYAU. (Hermann), Tome 1, p.157-164. Tome 2. p. 136-139.

(15) Frenkel, WAVE MECHANICS, p.317.

(16) " " " p.321-323.

(17) " " " p.318-319.

(18) L'ÉLECTRON MAGNÉTIQUE, p.140-142 and p.204-207.

(19) THEORETICAL PHYSICS, p.259.

(ref.20) L'ÉLECTRON MAGNÉTIQUE, p.286-289.
Frenkel: WAVE MECHANICS, p.346.

(21) L'ÉLECTRON MAGNÉTIQUE, chap. XX.

(22) M.Born: ATOMIC PHYSICS, (Blackie, 1944) p.184.
H.Wilson: MODERN PHYSICS, p.173.

(23) Sommerfeld: WAVE MECHANICS, chap.2, para. 10.
L'ÉLECTRON MAGNÉTIQUE, chap. XVII.

(24) " " chap. XVIII.

(25) H.T.Flint: WAVE MECHANICS, p.112.
Frenkel: WAVE MECHANICS, p.259 etc.

(26) THEORETICAL PHYSICS, p.69.

(27) H.T.Flint: Phil. Mag. April 1940, p.330
May 1940, p.417.