The quantum theory of the electron

by Dirac (117 S.610, 1928)

The new quantum mechanics, when applied to the problem of the structure of the atom with point charge electrons, does not give~~s~~ results in agreement with experiment. The discrepancies consist of "duplexity" phenomena; the observed number of stationary states for an electron in an atom being twice the number given by the theory. To meet the difficulty, Gousmidt and Uhlenbeck have introduced the idea of an electron with a spin angular momentum of half a quantum and a magnetic moment of one Bohr magneton. This model for the electron has been fitted into the new mechanics by Pauli,* and Darwin,† working with an equivalent theory, has shown that it gives results in agreement with experiment for hydrogen-like spectra to the first oder of accuracy.

The question remains as to why Nature should have chosen this particular model for the electron instead of being satisfied with the point-charge. One would like to find some incompleteness in the previous methods of applying quantum mechanics to the point-charge electron such that, when removed, the whole of the duplexity phenomena follow without arbitrary assumptions. In the present paper it is shown that this is the case, the incompleteness of the previous theories lying in their disagreement with relativity, or alternatively, with the general transformation theory of quantum mechanics. It appears that the simplest Hamiltonian for a point-charge electron ~~model, at least as a first approximation. The most important failure of the model, at A~~ satisfying the requirements of both relativity and the general transformation theory leads to an explanation of all duplexity phenomena without further assumption. All the same there is a great deal of truth in the spinning electron model, at least as a first approximation. The most important failure of the model seems to be that the magnitude of the resultant orbital angular momentum of an electron moving in an orbit in a central field of force is not a constant, as the model leads one to expect.

---

* Pauli 43 S.601

† Darwin 116 p.227

§1. Previous Relativity Treatments†

---

† Gordon, Klein.

This appears to be satisfactory so far as emission and absorption ~~is~~ of radiation are concerned, but is not so general as the interpretation of the non relativity quantum mechanics, which has been developed‡ sufficiently to enable one to answer the question: What is the probability of any dynamical variable at any specified time having a value lying between any specified limits, when the system is represented by a given wave function ψ_n? The Gordon–Klein interpretation can answer such questions if they refer to the position of the electron (by the use of ρ_{nn}), but not if they refer to its momentum, or angular momentum or any other dynamical variable. We should expect the interpretation of the relativity theory to be just as general as that of the non-relativity theory.

The general interpretation of non-relativity quantum mechanics is based on the transformation theory, and is made possible by the wave equation being of the form

$$(H - W)\psi = 0 \qquad (2)$$

i.e. being linear in W or $\frac{\partial}{\partial t}$, so that the wave function at any time determines the wave function at any later time. The wave equation of the relativity theory must also be linear in W if the general interpretation is to be possible.

The second difficulty in Gordon's interpretation arises from the fact that if one takes the conjugate imaginary of ~~the~~ equation (1), one gets

$$\left[\left(-\frac{W}{c} + \frac{e}{c}A_0\right)^2 + \left(-p + \frac{e}{c}A\right)^2 + m^2c^2\right]\psi = 0$$

which is the same as one would get if one put $-e$ for e. The wave equation (1) thus refers equally well to an electron with charge e as to one with charge $-e$. If one considers for definite~~d~~ness the limiting case of large quantum numbers one would find that some of the solutions of the wave equation are wave packets moving in the way a particle of charge $-e$ would move on the classical theory, while others are wave packets moving in the way a particle of charge e would move classically. For this second class of solutions W has a negative value. One gets over the difficulty on the classical theory by arbitrarily excluding those solutions that have a negative W. One cannot do this on the quantum theory, since in general a perturbation will cause transitions from states with W positive to states with W negative. Such a transitions ~~from states~~ would appear experimentally as the electron suddenly changing its charge ~~for~~ from $-e$ to e, a phenomenon which has not been observed. The true relativity wave equation should thus be such that its solutions split up into two non-combining sets, referring respectively to the charge $-e$ and the charge e.

In the present paper we shall be concerned only with the removal of the first of these two difficulties. The resulting theory is therefore still only an approximation, but it appears to be good enough to account for all the duplexity phenomena without arbitrary assumptions.

§2. The Hamiltonian for no Field.

Our problem is to obtain a wave equation of the form (2) which shall be invariant under a Lorentz transformation and shall be equivalent to (1) in the limit of large quantum numbers. We shall consider first the case of no field when equation (1) reduces to

$$(-p_0^2 + \mathbf{p}^2 + m^2c^2)\psi = 0 \qquad (3)$$

if one puts

$$p_0 = \frac{W}{c} = ih\frac{\partial}{\partial t}.$$

The symmetry between p_0 and p_1, p_2, p_3, required by the relativity shows that since the Hamiltonian we want is linear in p_0, it ~~is~~ must also be linear in p_1, p_2 and p_3. Our wave equation is therefore of the form

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\psi = 0, \qquad (4)$$

where for the present all that is known about the dynamical variables

~~of the~~ or operators $\alpha_1, \alpha_2, \alpha_3, \beta$ is ~~the~~ that they are independent of p_0, p_1, p_2, p_3, i.e. that they commute with t, x_1, x_2, x_3. Since we are considering the case of a particle moving in empty space, so that all points in space are equivalent, we should expect the Hamiltonian not to involve t, x_1, x_2, x_3. This means that $\alpha_1, \alpha_2, \alpha_3, \beta$ are independent of t, x_1, x_2, x_3, i.e. that they commutes with $p_0\ p_1\ p_2\ p_3$. We are therefore obliged to have other dynamical variables besides the co-ordinates and momenta of the electron, in order that $\alpha_1, \alpha_2, \alpha_3, \beta$ may be functions of them. The wave function ψ must then involve more variables than merely $x_1\ x_2\ x_3, t$.

Equation (4) leads to

$$\begin{aligned} 0 &= (-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\psi \\ &= [-p_0^2 + \Sigma\alpha_1^2 p_1^2 + \Sigma(\alpha_1\alpha_2 + \alpha_2\alpha_1)p_1p_2 + \beta^2 + \Sigma(\alpha_1\beta + \beta\alpha_1)p_1]\psi \end{aligned} \qquad (5)$$

where the Σ refers to cyclic permutation of the suffixes 1, 2, 3. This agrees with (3) if

$$\left.\begin{aligned} &\alpha_r^2 = 1 \qquad \alpha_r\alpha_s + \alpha_s\alpha_r = 0 \quad (r \neq s) \\ &\beta^2 = m^2c^2 \qquad \alpha_r\beta + \beta\alpha_r = 0 \end{aligned}\right\} \quad r, s = 1, 2, 3.$$

If we puts $\beta = \alpha_4 mc$, these conditions become

$$\alpha_\mu^2 = 1 \qquad \alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0 \quad (\mu \neq \nu) \quad \mu, \nu = 1, 2, 3, 4 \ (6)$$

We can suppose the α_μ's to be expressed as matrices in some matrix scheme, the matrix elements of α_μ being, say, $\alpha_\mu(\zeta'\zeta'')$. The wave function ψ must now be a function of ζ as well as x_1, x_2, x_3, t. The result of α_μ multiplied into ψ will be a function $(\alpha_\mu\psi)$ of x_1, x_2, x_3, t, ζ defined by

$$(\alpha_\mu\psi)(x, t, \zeta) = \Sigma_{\zeta'}\,\alpha_\mu(\zeta\zeta')\psi(x, t, \zeta')$$

We must now find four matrices α_μ to satisfy the conditions (6). We make use of the matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which Pauli* introduced to describe the three components of spin angular momentum. These matrices have just the properties

$$\sigma_r^2 = 1 \qquad \sigma_r\sigma_s + \sigma_s\sigma_r = 0 \qquad (r \neq s) \qquad (7)$$

That we require for our α's. We cannot, however, just take the σ's to be three of our α's, because then it would not be possible to find the fourth. We must extend the σ's in a diagonal manner to bring in two more rows and columns, so that we can introduce three more matrices ρ_1, ρ_2, ρ_3 of the same form as $\sigma_1, \sigma_2, \sigma_3$, but referring to different rows and columns, thus :—

$$\sigma_1 = \begin{Bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{Bmatrix} \qquad \sigma_2 = \begin{Bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{Bmatrix} \qquad \sigma_3 = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}$$

$$\rho_1 = \begin{Bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{Bmatrix} \qquad \rho_2 = \begin{Bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{Bmatrix} \qquad \rho_3 = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}$$

The ρ's are obtained from the σ's by interchanging the second and third rows, and the second and third columns. We now have, in addition to equations (7)

$$\left.\begin{aligned} \rho_r^2 = 1 . \qquad \rho_r\rho_s + \rho_s\rho_r &= 0 \qquad (r \neq s) \\ \rho_r\sigma_t &= \sigma_t\rho_r \end{aligned}\right\} \qquad (7')$$

and also
If we now take

$$\alpha_1 = \rho_1\sigma_1, \quad \alpha_2 = \rho_1\sigma_2, \quad \alpha_3 = \rho_1\sigma_3, \quad \alpha_4 = \rho_3$$

all the conditions (6) are satisfied, e.g.

$$\alpha_1^2 = \rho_1\sigma_1\rho_1\sigma_1 = \rho_1^2\sigma_1^2 = 1$$

$$\alpha_1\alpha_2 = \rho_1\sigma_1\rho_1\sigma_2 = \rho_1^2\sigma_1\sigma_2 = -\rho_1^2\sigma_2\sigma_1 = -\alpha_2\alpha_1$$

The following equations are to be noted for later reference

$$\left.\begin{aligned} \rho_1\rho_2 &= i\rho_3 = -\rho_2\rho_1 \\ \sigma_1\sigma_2 &= i\sigma_3 = -\sigma_2\sigma_1 \end{aligned}\right\} \qquad (8)$$

$$\cancel{(\rho_2+\rho_1)\rho_2 = \rho_3(\rho_2+\rho_3)}$$

---

*Pauli, loc. cit.

$$x^* = \frac{x'-vt'}{\sqrt{1-\frac{v^2}{c^2}}} = \kappa(x'-vt') \qquad y = y' \qquad z = z' \qquad t = \kappa\left(t' + \frac{vx'}{c^2}\right)$$

$$dx^2 = \kappa^2(dx - vdt)^2$$

$$-c^2dt^2 = (c^2\kappa^2)\left(dt^* + \frac{v}{c^2}dx\right)^2$$

$$\frac{\partial}{\partial x'} = \kappa\frac{\partial}{\partial x} + \frac{\kappa v}{c^2}\frac{\partial}{\partial t} \qquad \frac{\partial}{\partial y'} = \frac{\partial}{\partial y} \qquad \frac{\partial}{\partial z'} = \frac{\partial}{\partial z} \qquad \frac{1}{c}\frac{\partial}{\partial t'} = -\frac{\kappa v}{c}\frac{\partial}{\partial x^*} + \frac{\kappa}{c}\frac{\partial}{\partial t^*}$$

together with the equations obtained by cyclic permutation of the suffixes.

The wave equation (4) now take the form

$$[p_0 + \rho_1(\sigma, p) + \rho_3 mc]\psi = 0, \qquad (9)$$

where σ denotes the vector $(\sigma_1\ \sigma_2\ \sigma_3)$.

§3. Proof of Invariance under a Lorentz Transformation

Multiply equation (9) by ρ_3 on the left-hand side. It becomes, with the help of (8),

$$[\rho_3 p_0 + i\rho_2(\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3) + mc]\psi = 0.$$

Putting

$$p_0 = ip_4,$$

$$\rho_3 = \gamma_4, \qquad \rho_2\sigma_r = \gamma_r, \qquad r = 1, 2, 3, \qquad (10)$$

we have

$$[i\Sigma\gamma_\mu p_\mu + mc]\psi = 0, \qquad \mu = 1, 2, 3, 4. \qquad (11)$$

The p_μ transform under a Lorentz transformation according to the law

$$p_\mu' = \Sigma_\nu a_{\mu\nu} p_\nu, \qquad \frac{\partial}{\partial x} = a_{\mu\nu} \qquad \begin{Bmatrix} \kappa & 0 & 0 & +\frac{\kappa v}{c} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{\kappa v}{c} & 0 & 0 & \kappa \end{Bmatrix}$$

where the coefficients $a_{\mu\nu}$ are c-numbers satisfying

$$\Sigma_\mu a_{\mu\nu}a_{\mu\tau} = \delta_{\nu\tau}, \qquad \Sigma_\tau a_{\mu\tau}a_{\nu\tau} = \delta_{\mu\nu}.$$

$$\kappa^2(1-v^2)$$

The wave equation therefore transforms into

$$[i\Sigma\gamma_\mu' p_\mu' + mc]\psi = 0, \qquad (12)$$

where

$$\gamma_\mu' = \Sigma_\nu a_{\mu\nu}\gamma_\nu.$$

Now the γ_μ, like the α_μ, satisfy

$$\gamma_\mu^2 = 1, \qquad \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 0 \qquad (\mu \neq \nu).$$

These relations can be summed up in the single equation

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}.$$

We have

$$\begin{aligned} \gamma_\mu'\gamma_\nu' + \gamma_\nu'\gamma_\mu' &= \Sigma_{\tau\lambda} a_{\mu\tau}a_{\nu\lambda}(\gamma_\tau\gamma_\lambda + \gamma_\lambda\gamma_\tau) \\ &= 2\Sigma_{\tau\lambda} a_{\mu\tau}a_{\nu\tau}\delta_{\tau\lambda} \\ &= 2\Sigma_\tau a_{\mu\tau}a_{\nu\tau} = 2\delta_{\mu\nu}. \end{aligned}$$

Thus the γ_μ' satisfy the same relations as the γ_μ. Thus we can put, analogously to (10)

$$\gamma_4' = \rho_3' \qquad \gamma_r' = \rho_2'\sigma_r'$$

where the ρ's and σ's are easily verified to satisfy the relations cor-

-responding to (7), (7') and (8), if ρ_2' and ρ_1' are defined by $\rho_2' = -i\gamma_1'\gamma_2'\gamma_3'$, $\rho_1' = -i\rho_2'\rho_3'$.

We shall now show that, by a canonical transformation, the ρ's and σ's may be brought into the form of the ρ's and σ's. From the equation $\rho_3'^2 = 1$, it follows that the only possible characteristics value for ρ_3' are ± 1. If one applies to ρ_3' a canonical transformation with the transformation function ρ_1', the result is

$$\rho_1'\rho_3'(\rho_1')^{-1} = -\rho_3'\rho_1'(\rho_1')^{-1} = -\rho_3'.$$

Since characteristic values are not changed by a ~~con~~ canonical transformation, ρ_3' must have the same characteristic values as $-\rho_3'$. Hence the characteristic values of ρ_3' are $+1$ twice and -1 twice. The same argument applies to each of the other ρ's, and to each of the σ's.

Since ρ_3' and σ_3' commute, they can be brought simultaneously to the diagonal form by a canonical transformation. They will then have for their diagonal elements each $+1$ twice and -1 twice. Thus by suitably rearranging the rows and columns, they can be brought into the form ρ_3 and σ_3 respectively. (The possibility $\rho_3' = \pm\sigma_3'$ is excluded by the existence of matrices that commute with one but not with the other.) Any matrix containing four rows and columns can be expressed as

$$c + \Sigma_r c_r\sigma_r + \Sigma_r c_r'\rho_r + \Sigma_{rs} c_{rs}\rho_r\sigma_s \qquad (13)$$

where the sixteen coefficients c, c_r, c_r', c_{rs} are c-numbers. By expressing σ_1' in this way we seen, from the fact that it commutes with $\rho_3' = \rho_3$ and anticommutes* with $\sigma_3' = \sigma_3$, that it must be of the form

$$\sigma_1' = c_1\sigma_1 + c_2\sigma_2 + c_{31}\rho_3\sigma_1 + c_{32}\mp\rho_3\sigma_2,$$

i.e. of the form

$$\sigma_1' = \begin{Bmatrix} 0 & a_{12} & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{Bmatrix}$$

The ~~equat~~ condition $\sigma_1'^2 = 1$ shows that $a_{12}a_{21} = 1$, and, $a_{34}a_{43} = 1$. If we now apply the canonical transformation: first row to be multiplied by $(a_{21}/a_{12})^{\frac{1}{2}}$ and third row to be multiplied by $(a_{43}/a_{34})^{1/2}$, and ~~that~~ first and third columns to be divided by the same expressions, σ_1' will be brought into the form of σ_1, and the diagonal matrices σ_3' and ρ_3' will not be changed.

If we now express ρ_1' in the form (13) and use the conditions that it commutes with $\sigma_1' = \sigma_1$ and $\sigma_3' = \sigma_3$ and anticommutes with $\rho_3' = \rho_3$, we see that it must be ~~co~~ of the form

$$\rho_1' = c_1'\rho_1 + c_2'\rho_2.$$

The condition $\rho_1'^2 = 1$ shows that $c_1'^2 + c_2'^2 = 1$, or $c_1' = \cos\theta$, $c_2' = \sin\theta$. Hence ρ_1' is of the form

$$\rho_1' = \begin{Bmatrix} 0 & 0 & e^{-i\theta} & 0 \\ 0 & 0 & 0 & e^{-i\theta} \\ e^{i\theta} & 0 & 0 & 0 \\ 0 & e^{i\theta} & 0 & 0 \end{Bmatrix}$$

If we apply the canonical transformation~~s~~; first and second rows to be multipled by $e^{i\theta}$ and first and second columns to be divided by the same expression, ρ_2' will be of the form ρ_2 and σ_2, on account of the relations $i\rho_2' = \rho_3'\rho_1'$, $i\sigma_2' = \sigma_3'\sigma_1'$.

Thus by a succession of canonical transformations, which can be combined to form a single canonical transformation, the ρ's and σ's can be brought into the form of the original wave equation (11) or (9), so that the results that follow~~s~~ from this original wave equation must be independent of the frame of reference used.

§ The Hamiltonian for an arbitrary field.

To obtain the Hamiltonian for an electron in an electromagnetic field with ~~the~~ scalar potential A_0 and vector potential A, we adopt the usual procedure of substituting $p_0 + e/c\,A_0$ for p_0 and $p + e/c\,A$ for p in the Hamiltonian for no field. From equation (9) we thus obtain

$$[p_0 + \frac{e}{c}A_0 + \cancel{\rho_3}\,\rho_1(\sigma, p + \frac{e}{c}A) + \rho_3 mc]\psi = 0. \qquad (14)$$

This wave equation~~s~~ appears to be sufficient to account for all the duplexity phenomena. On account of the matrices ρ and σ containing four rows and columns, it will have four times as many solutions as the non-relativity wave equation, and twice as may as the previous relativity wave equation (1). Since half the solutions must be rejected as referring to the charge $+e$ on the electron, the correct number will be left to account for duplexity phenomena. The proof given in the preceding ~~set~~ section of invariance under a Lorentz transformation applies equally well to the more general wave equation (14).

We can obtain a rough idea of how (14) differs from the previous relativity wave equation (1) by multiplying it up analogously to (5). This gives, if we write e' for e/c

$$
\begin{aligned}
0 = & \left[-(p_0+e'A_0)+\rho_1(\boldsymbol{\sigma}, \mathbf{p}+e'\mathbf{A})+\rho_3 mc\right] \\
& \times\left[(p_0+e'A_0)+\rho_1(\boldsymbol{\sigma}, \mathbf{p}+e'\mathbf{A})+\rho_3 mc\right]\psi \\
= & \left[-(p_0+e'A_0)^2+(\boldsymbol{\sigma}, \mathbf{p}+e\mathbf{A})^2+m^2c^2\right. \\
& \left.+\rho_1\{(\boldsymbol{\sigma}, \mathbf{p}+e'\mathbf{A})(p_0+e'A_0)-(p_0+e'A_0)(\boldsymbol{\sigma}_0, \mathbf{p}+e'\mathbf{A})\}\right]\psi
\end{aligned}
\tag{15}
$$

We now use the general formula, that if $\mathbf{B}$ and $\mathbf{C}$ are any two vectors that commute with $\boldsymbol{\sigma}$

$$
\begin{aligned}
(\boldsymbol{\sigma}, \mathbf{B})(\boldsymbol{\sigma}, \mathbf{C}) &= \Sigma\sigma_1^2 B_1C_1+\Sigma(\sigma_1\sigma_2 B_1C_2+\sigma_2\sigma_1 B_2C_1) \\
&= (\mathbf{B}, \mathbf{C})+i\Sigma\sigma_3(B_1C_2-B_2C_1) \\
&= (\mathbf{B}, \mathbf{C})+i(\boldsymbol{\sigma}, \mathbf{B}\times\mathbf{C}).
\end{aligned}
\tag{16}
$$

Taking $\mathbf{B}=\mathbf{C}=\mathbf{p}+e'\mathbf{A}$, we find

$$
\begin{aligned}
(\boldsymbol{\sigma}, \mathbf{p}+e'\mathbf{A})^2 &= (\mathbf{p}+e'\mathbf{A})^2+i\Sigma\sigma_3\left[(p_1+e'A_1)(p_2+eA_2)-(p_2+e'A_2)(p_1+e'A_1)\right] \\
&= (\mathbf{p}+e'\mathbf{A})^2+he'(\boldsymbol{\sigma}, \operatorname{curl}\mathbf{A})
\end{aligned}
$$

Thus (15) becomes

$$
\begin{aligned}
0 = & \left[-(p_0+e'A_0)^2+(\mathbf{p}+e'\mathbf{A})^2+m^2c^2+e'h(\boldsymbol{\sigma}, \operatorname{curl}\mathbf{A})\right. \\
& \left.-ie'h\rho_1\left(\boldsymbol{\sigma}, \operatorname{grad}A_0+\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}\right)\right]\psi \\
= & -(p_0+e'A_0)^2+(\mathbf{p}+e'\mathbf{A})^2+m^2c^2+e'h(\boldsymbol{\sigma}, \mathbf{H})+ie'h\rho_1(\boldsymbol{\sigma}, \mathbf{E})]\psi,
\end{aligned}
$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic vector of the field. This differs from (1) by the two extra terms

$$\frac{eh}{c}(\boldsymbol{\sigma}, \mathbf{H})+\frac{ieh}{c}\rho_1(\boldsymbol{\sigma}, \mathbf{E})$$

in F. These two terms, when divided by the factor $2m$, can be regarded as the additional potential energy of the electron due to its new degree of freedom. The electron will therefore ~~be~~ behave as though it has a magnetic moment $eh/2mc\cdot\boldsymbol{\sigma}$ and an electric moment $ieh/2mc\cdot\rho_1\boldsymbol{\sigma}$. This magnetic moment is just that assumed in the spinning electron model. The electric moment, being a pure imaginary, we should not expect to appear in the model. It is doubtful whether the electric moment has any physical meaning, since the Hamiltonian in (14) that we started from the real, and the imaginary part only appeared when we multiplied it up in an artificial way in order to make it resemble the Hamiltonian of previous theories.

§ The Angular Momentum Integrals for Motion in a Central Field.

We shall consider in greater detail the motion of an electron in a central field of force. We put $\mathbf{A}=0$ and $e'A_0=V(r)$, an arbitrary function of the radius r, so that the Hamiltonian in (14) becomes

$$F\equiv p_0+V+\rho_1(\boldsymbol{\sigma}, \mathbf{p})+\rho_3 mc.$$

We shall determine the periodic solutions of the wave equation $F\psi=0$, which means that p_0 is to be counted as a parameter instead of an operator; it is, in fact, just $1/c$ times the energy level.

We shall first find the angular momentum integrals of the motion. The orbital angular momentum $\mathbf{m}$ is defined by

$$\mathbf{m}=\mathbf{x}*\mathbf{p},$$

and satisfies the following "Vertauschungs" relations

$$
\left.
\begin{array}{ll}
m_1x_1-x_1m_1=0, & m_1x_2-x_2m_1=ihx_3 \\
m_1p_1-p_1m_1=0, & m_1p_2-p_2m_1=ihp_3 \\
\mathbf{m}\times\mathbf{m}=ih\mathbf{m} & \mathbf{m}^2m_1-m_1\mathbf{m}^2=0
\end{array}
\right\}
\tag{17}
$$

Together with similar relations obtained by permuting the suffixes. Also we

commutes with r, and with p_r, the momentum canonically conjugate to r.
We have

$$\begin{aligned} m_1 F - F m_1 &= \rho_1 \{ m_1 (\boldsymbol{\sigma}, \mathbf{p}) - (\boldsymbol{\sigma}, \mathbf{p}) m_1 \} \\ &= \rho_1 (\boldsymbol{\sigma}, m_1 \mathbf{p} - \mathbf{p} m_1) \\ &= i h \rho_1 (\sigma_2 p_3 - \sigma_3 p_2), \end{aligned}$$

and so

$$\mathbf{m} F - F \mathbf{m} = i h \rho_1 \boldsymbol{\sigma} \times \mathbf{p} \qquad (18)$$

Thus $\mathbf{m}$ is not a constant of the motion. We have further

$$\begin{aligned} \sigma_1 F - F \sigma_1 &= \rho_1 \{ \sigma_1 (\boldsymbol{\sigma}, \mathbf{p}) - (\boldsymbol{\sigma}, \mathbf{p}) \sigma_1 \} \\ &= \rho_1 (\sigma_1 \boldsymbol{\sigma} - \boldsymbol{\sigma} \sigma_1, \mathbf{p}) \\ &= 2 i \rho_1 (\sigma_3 p_2 - \sigma_2 p_3), \end{aligned}$$

with the help of (8), and so

$$\boldsymbol{\sigma} F - F \boldsymbol{\sigma} = -2 i \rho_1 \boldsymbol{\sigma} \times \mathbf{p}.$$

Hence

$$(\mathbf{m} + \tfrac{1}{2} h \boldsymbol{\sigma}) F - F (\mathbf{m} + \tfrac{1}{2} h \boldsymbol{\sigma}) = 0.$$

Thus $\mathbf{m} + \frac{1}{2} h \boldsymbol{\sigma}$ (= $\mathbf{M}$ say) is a constant of the motion. We can interpret this result by saying that the electron has a spin angular momentum of $\frac{1}{2} h \boldsymbol{\sigma}$, which, added to the orbital angular momentum $\mathbf{m}$, gives the total angular momentum $\mathbf{M}$, which is a constant of the motion.

The Vertauschungs relations (17) all hold when M's are written for the m's. In particular

$$\mathbf{M} \times \mathbf{M} = i h \mathbf{M} \quad \text{and} \quad \mathbf{M}^2 M_3 = M_3 \mathbf{M}^2.$$

M_3 will be an action variable of the system. ~~Sys~~ Since the characteristic values of m_3 must be integral multiples of h in ~~od~~ order that the wave function may be single-valued, the characteristic values of M_3 must be half odd integral multiples of h. If we put

$$\mathbf{M}^2 = (j^2 - \tfrac{1}{4}) h^2, \qquad (19)$$

j will be another quantum number, and the characteristic values of M_3 will extend from $(j - \frac{1}{2})h$ to $(-j + \frac{1}{2})h$.* Thus j takes integral values.

One easily verifies from (18) that $\mathbf{m}^2$ does not commute with F, and is thus not a constant of the motion. This makes a difference between the present theory and the previous spinning electron theory, in which $\mathbf{m}^2$ is constant,

---

* see III. p. 281

$(\psi_\alpha \varphi_\beta)$
$(\sigma \psi)$

$W \rightarrow$ ~~$\Psi(\psi_\alpha \psi_\beta \psi_\gamma \psi_\delta)$~~

and defines the azimutal quantum number k by a relation similar to (19). We shall find that our j plays the same part as the k of the previous theory.

$\varphi_1 \psi_1, \varphi_1 \psi_2, \ldots$

§6. The Energy Levels for Motion in a Central Field.

We shall now obtain the wave equation as a differential equation in r, with the variables that specify the orientation of the whole system removed. We can do this by the use only of elementary non-commutative algebra in the following way.

In formula (16) take $B = C = \mathbf{m}$. This gives

$$\begin{aligned} (\boldsymbol{\sigma}, \mathbf{m})^2 &= \mathbf{m}^2 + i (\boldsymbol{\sigma}, \mathbf{m} \times \mathbf{m}) \\ &= (\mathbf{m} + \tfrac{1}{2} h \boldsymbol{\sigma})^2 - h (\boldsymbol{\sigma}, \mathbf{m}) - \tfrac{1}{4} h^2 \boldsymbol{\sigma}^2 - h (\boldsymbol{\sigma}, \mathbf{m}) \\ &= \mathbf{M}^2 - 2h (\boldsymbol{\sigma}, \mathbf{m}) - \tfrac{3}{4} h^2. \end{aligned} \qquad (20)$$

Hence

$$\{ (\boldsymbol{\sigma}, \mathbf{m}) + h \}^2 = \mathbf{M}^2 + \tfrac{1}{4} h^2 = j^2 h^2$$

Up to the present we have defined j only through j^2, so that we could now, if we liked, take jh equal to $(\boldsymbol{\sigma}, \mathbf{m}) + h$. This would not be convenient since we want j to be a constant of the motion while $(\boldsymbol{\sigma}, \mathbf{m}) + h$ is not, although its square is. We have, in fact, by another application of (16)

$$(\boldsymbol{\sigma}, \mathbf{m})(\boldsymbol{\sigma}, \mathbf{p}) = i (\boldsymbol{\sigma}, \mathbf{m} \times \mathbf{p})$$

since $(\mathbf{m}, \mathbf{p}) = 0$, and similarly

$$(\boldsymbol{\sigma}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{m}) = i (\boldsymbol{\sigma}, \mathbf{p} \times \mathbf{m}),$$

so that

$$\begin{aligned} (\boldsymbol{\sigma}, \mathbf{m})(\boldsymbol{\sigma}, \mathbf{p}) + (\boldsymbol{\sigma}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{m}) &= i \Sigma \sigma_1 (m_2 p_3 - m_3 p_2 + p_2 m_3 - p_3 m_2) \\ &= i \Sigma \sigma_1 \cdot 2 i h p_1 = -2h (\boldsymbol{\sigma}, \mathbf{p}). \end{aligned}$$

or

$$\{ (\boldsymbol{\sigma}, \mathbf{m}) + h \} (\boldsymbol{\sigma}, \mathbf{p}) + (\boldsymbol{\sigma}, \mathbf{p}) \{ (\boldsymbol{\sigma}, \mathbf{m}) + h \} = 0.$$

Thus $(\boldsymbol{\sigma}, \mathbf{m}) + h$ anticommutes with one of the terms in F, namely, $\rho_1 \times (\boldsymbol{\sigma}, \mathbf{p})$, and commutes with the other three. Hence $\rho_3 \{ (\boldsymbol{\sigma}, \mathbf{m}) + h \}$ commutes with all four, and is therefore a constant of the motion. But the square of $\rho_3 \{ (\boldsymbol{\sigma}, \mathbf{m}) + h \}$ must also equal $j^2 h^2$. We therefore take

$$\boxed{j} h = \rho_3 \{ (\boldsymbol{\sigma}, \mathbf{m}) + h \}. \qquad (21)$$

We have, by a further application of (16)

$$(\sigma, \mathbf{x})(\sigma, \mathbf{p}) = (\mathbf{x}, \mathbf{p}) + i(\sigma, \mathbf{m}).$$

Now a permissible definition of p_r is

$$(\mathbf{x}, \mathbf{p}) = r p_r + ih, ?$$

and from (21)

$$(\sigma, \mathbf{m}) = \rho_3 j h - h.$$

Hence $$(\sigma, \mathbf{x})(\sigma, \mathbf{p}) = r p_r + i\rho_3 j h. \qquad (22)$$

Introduce the quantity ε defined by

$$r\varepsilon = \rho_1(\sigma, \mathbf{x}) \qquad (23)$$

Since r commutes with ρ_1 and with $(\sigma, \mathbf{x})$, it must commute with ε. We thus have

$$r^2\varepsilon^2 = [\rho(\sigma, \mathbf{x})]^2 = (\sigma, \mathbf{x})^2 = \mathbf{x}^2 = r^2$$

or $$\varepsilon^2 = 1$$

Since there is symmetry between $\mathbf{x}$ and $\mathbf{p}$ so far as angular momentum is concerned, $\rho_1(\sigma, \mathbf{x})$, like $\rho(\sigma, \mathbf{p})$, must commute with $\mathbf{M}$ and j. Hence ε commutes with $\mathbf{M}$ and j. Further, ε must commute with p_r, since we have

$$(\sigma, \mathbf{x})(\mathbf{x}, \mathbf{p}) - (\mathbf{x}, \mathbf{p})(\sigma, \mathbf{x}) = ih(\sigma, \mathbf{x}),$$

which gives $$r\varepsilon(rp_r + ih) - (rp_r + ih)r\varepsilon = ihr\varepsilon,$$

which reduces to $$\varepsilon p_r - p_r\varepsilon = 0.$$

From (22) and (23) we now have

$$r\varepsilon\rho_1(\sigma, \mathbf{p}) = rp_r + i\rho_3 j h$$

or $$\rho_1(\sigma, \mathbf{p}) = \varepsilon p_r + i\varepsilon\rho_3 jh/r.$$

Thus $$F = p_0 + V + \varepsilon p_r + i\varepsilon\rho_3 jh/r + \rho_3 mc. \qquad (24)$$

Equation (23) shows that ε anticommutes with ρ_3. We can therefore by a canonical transformation (involving perhaps the x's and p's as well as the σ's and ρ's) bring ε into the form of ~~ρ.~~ the ρ_2 of § 2 without changing ρ_3, and without changing any of the other variables occurring on the right-hand side of (24), since these other variables all commute with ε. $i\varepsilon\rho_3$ will now be of the form $i\rho_2\rho_3 = -\rho_1$, so that the wave equation takes the form.

$$F\psi \equiv (p_0 + V + \rho_2 p_r - \rho_1 jh/r + \rho_3 mc)\psi = 0.$$

$$\rho_1\varepsilon = \varepsilon\rho_1 \qquad \rho_1\rho_2 = -\rho_2\rho_1$$
$$\rho_3\varepsilon = -\varepsilon\rho_3 \qquad \rho_3\rho_2 = -\rho_2\rho_3.$$
$$T\varepsilon = \rho_2 T.$$

If we write this equations out in full, calling the components of ψ referring to the first and third rows (or columns) of the matrices ψ_α and ψ_β respectively we get

$$(F\psi)_\alpha \equiv (p_0 + V)\psi_\alpha - h\frac{\partial}{\partial r}\psi_\beta - \frac{jh}{r}\psi_\beta + mc\psi_\alpha = 0,$$
$$(F\psi)_\beta \equiv (p_0 + V)\psi_\beta + h\frac{\partial}{\partial r}\psi_\alpha - \frac{jh}{r}\psi_\alpha - mc\psi_\beta = 0.$$

The second and fourth components gives just a repetition of these two equations. We shall now eliminate ψ_α. If we write hB for $p_0 + V + mc$, the first equation becomes

$$\left(\frac{\partial}{\partial r} + \frac{j}{r}\right)\psi_\beta = B\psi_\alpha$$

which gives on differentiating

$$\frac{\partial^2}{\partial r^2}\psi_\beta + \frac{j}{r}\frac{\partial}{\partial r}\psi_\beta - \frac{j}{r^2}\psi_\beta = B\frac{\partial}{\partial r}\psi_\alpha + \frac{\partial B}{\partial r}\psi_\alpha$$
$$= \frac{B}{h}\left[-(p_0 + V - mc)\psi_\beta + \frac{jh}{r}\psi_\alpha\right] + \frac{1}{h}\frac{\partial V}{\partial r}\psi_\alpha$$
$$= -\frac{(p_0 + V)^2 - m^2c^2}{h^2}\psi_\beta + \left(\frac{j}{r} + \frac{1}{Bh}\frac{\partial V}{\partial r}\right)\left(\frac{\partial}{\partial r} + \frac{j}{r}\right)\psi_\beta$$

This reduces to

$$\frac{\partial^2}{\partial r^2}\psi_\beta + \left[\frac{(p_0 + V)^2 - m^2c^2}{h^2} - \frac{j(j+1)}{r^2}\right]\psi_\beta - \frac{1}{Bh}\frac{\partial V}{\partial r}\left(\frac{\partial}{\partial r} + \frac{j}{r}\right)\psi_\beta = 0. \qquad (25)$$

The values of the parameter p_0 for which this equation has a solution finite at $r = 0$ and $r = \infty$ are $1/c$ times the energy levels of the system. To compare this equation with those of previous theories, we put $\psi_\beta = r\chi$, so that

$$\frac{\partial^2}{\partial r^2}\chi + \frac{2}{r}\frac{\partial}{\partial r}\chi + \left[\frac{(p_0 + V)^2 - m^2c^2}{h^2} - \frac{j(j+1)}{r^2}\right]\chi$$
$$- \frac{1}{Bh}\frac{\partial V}{\partial r}\left(\frac{\partial}{\partial r} + \frac{j+1}{r}\right)\chi = 0 \qquad (26)$$

If one neglects the last term which is small on account of B being large, this equation becomes the same as the ordinary Schroedinger equation for the system, with relativity correction included. Since j has, from its definition, both positive and negative integral characteristic values, our equation will give twice as many energy levels when the last term is not neglected.

We shall now compare the last term of (26), which is of the same order

$$p_0 = \left(mc + \frac{W}{c}\right)$$

of magnitude as the relativity correction, with the spin correction given by Darwin and Pauli. To do this we must eliminate the $\partial\chi/\partial r$ term by a further transformation of the wave function. We put

$$\chi = B^{-\frac{1}{2}}\cdot\chi_1,$$

which gives

$$\frac{\partial^2}{\partial r^2}\chi_1 + \frac{2}{r}\frac{\partial}{\partial r}\chi_1 + \left[\frac{(p_0+V)^2 - m^2c^2}{h^2} - \frac{j(j+1)}{r^2}\right]\chi_1$$
$$+ \left[\frac{1}{Bh}\frac{j}{r}\frac{\partial V}{\partial r} - \frac{1}{2}\frac{1}{Bh}\frac{\partial^2 V}{\partial r^2} + \frac{1}{4}\frac{1}{B^2h^2}\left(\frac{\partial V}{\partial r}\right)^2\right]\chi_1 = 0 \quad (27)$$

The correction is now, to the 1st order of accuracy

$$\frac{1}{Bh}\left(\frac{j}{r}\frac{\partial V}{\partial r} - \frac{1}{2}\frac{\partial^2 V}{\partial r^2}\right),$$

where $Bh = 2mc$ (provided p_0 is positive). For the hydrogen atom we must put $V = e^2/cr$. The first order correction now becomes

$$-\frac{e^2}{2mc^2r^3}(j+1). \quad + \quad (28)$$
$$-$$

If we write $-j$ for $j+1$ in (27), we ~~do~~ do not alter the terms representing the unperturbed system, so

$$\frac{e^2}{2mc^2r^3}j \quad \pm \quad (28')$$

will give a second possible correction for the same unperturbed term.

In the theory of Pauli and Darwin, the corresponding correcting term is

$$\frac{e^2}{2mhc^2r^3}(\boldsymbol{\sigma}, \mathfrak{m})$$

when the Thomas factor $\frac{1}{2}$ is included. We must remember that in the Pauli-Darwin theory, the resultant orbital angular momentum k plays the part of our j. We must define k by

$$\mathfrak{m}^2 = k(k+1)h^2$$

instead of by the exact analogue of (19), in order that it may ~~be~~ have integral characteristic values, like j. We have from (20)

$$(\boldsymbol{\sigma}, \mathfrak{m})^2 = k(k+1)h^2 - h(\boldsymbol{\sigma}, \mathfrak{m})$$

or

$$\left\{(\boldsymbol{\sigma}, \mathfrak{m}) + \frac{1}{2}h\right\}^2 = \left(k+\frac{1}{2}\right)^2 h^2,$$

hence

$$(\boldsymbol{\sigma}, \mathfrak{m}) = kh \text{ or } -(k+1)h.$$

The correction becomes

$$\frac{e^2}{2mc^2r^3}k \quad \text{or} \quad -\frac{e^2}{2mc^2r^3}(k+1)$$

which agrees with (28) and (28)'. The present theory will thus, in the 1st approximation, leads to the same energy levels as those obtained by Darwin, which are in agreement with experiment.

The Quantum Theory of the Electron. Part II,
Dirac.

118, 351

In a previous paper by the author it is shown that the general theory of quantum mechanics together with relativity require the wave equation for an electron moving in an arbitrary electromagnetic field of potentials, A_0, A_1, A_2, A_3 to be of the form

$$F\psi \equiv [p_0 + \tfrac{e}{c}A_0 + \alpha_1(p_1 + \tfrac{e}{c}A_1) + \alpha_2(p_2 + \tfrac{e}{c}A_2) + \alpha_3(p_3 + \tfrac{e}{c}A_3) + \alpha_4 mc]\psi = 0 \qquad (1)$$

The α's are now dynamical variables which it is necessary to introduce in order to satisfy the conditions of the problem. They may be regarded as describing some internal motion of the electron, which for most purpose may be taken to be the spin of the electron postulated in previous theories. We shall call them the spin variables.

The α's must satisfy the conditions

$$\alpha_\mu^2 = 1 \qquad \alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0. \qquad (\mu \neq \nu)$$

They may conveniently be expressed in terms of six variables $\rho_1, \rho_2, \rho_3, \sigma_1, \sigma_2, \sigma_3$ that satisfy

$$\left.\begin{array}{ll} \rho_r^2 = 1, \quad \sigma_r^2 = 1, & \rho_r\sigma_s = \sigma_s\rho_r \qquad (r, s = 1, 2, 3) \\ \rho_1\rho_2 = i\rho_3 = -\rho_2\rho_1 & \sigma_1\sigma_2 = i\sigma_3 = -\sigma_2\sigma_1 \end{array}\right\}, (2)$$

together with the relations obtained from these by cyclic permutation of the suffixes, by means of the equations

$$\alpha_1 = \rho_1\sigma_1, \quad \alpha_2 = \rho_1\sigma_2, \quad \alpha_3 = \rho_1\sigma_3, \quad \alpha_4 = \rho_3,$$

The variables $\sigma_1, \sigma_2, \sigma_3$ now form the three components of a vector, which corresponds (apart from ~~the~~ a constant factor) to the spin angular momentum vector that appears in Pauli's theory of the spinning electron. The ρ's and σ's vary with the time, like other dynamical variables. Their equations of motion, written in the Poisson Bracket notation [ ], are

$$\dot{\rho}_r = c[\rho_r, F] \qquad \dot{\sigma} = c[\sigma_r, F].$$

It should be observed that these equations of motion are consistent with the conditions (2), so that if the conditions are satisfied initially they always remain satisfied. For example, we have

$$ih/c \cdot \dot{\sigma}_1 = \sigma_1 F - F\sigma_1 = 2i\rho_1\sigma_3(p_2 + e/c \cdot A_2) - 2i\rho_1\sigma_2(p_3 + e/c\, A_3).$$

Thus $\dot{\sigma}_1$ anticommutes with σ_1, so that

$$d\sigma_1^2/dt = \dot{\sigma}_1\sigma_1 + \sigma_1\dot{\sigma}_1 = 0.$$

The ρ's and σ's, and therefore also any function of them, can be represented by matrices with four rows and columns. A possible representation, in which ρ_3 and σ_3 are digonal matrices, is given in (loc. cit.) §2. Such a representation can apply only to a single instant of time, since the ρ's and σ's vary with the time. To get a scheme of representation which holds for all times, so that the equations of motion are valid in it, we should have to have only constants of the motion as diagonal matrices. It is, however, quite correct for the purpose of solving the wave equation (1) take a matrix representation for the ρ's and σ's which holds only for a single instant of time (as was done in loc. cit.); since the wave function is then the transformation function connecting the ρ's and σ's and x's at this particular time with a set of variables that are constants of the motion, as is required for the general interpretation of quantum mechanics.

Before we proceed with the theory of atoms with single electrons that was begun in loc. cit., the proof will be given of the conservation theorem, which states that the change in the probability of the electron being in a given volume during a given time is equal to ~~p~~ the probability of its having crossed the boundary. This proof is supplementary to the work of loc. cit. §3, and is necessary before one can infer that the theory will give consistent results that are invariant under a Lorentz transformation.

§ 1. The conservation Theorem

We shall first make a slight generalization of the usual interpretation of wave mechanics to apply to cases when the Hamiltonian is not Hermitian. Let the wave equation, written in certain variables q, be

$$(H - W)\psi = 0 \qquad (i)$$

Consider also the equation

$$(\hat{H} - \tilde{W})\phi = 0$$

or

$$(\tilde{H} + W)\phi = 0, \qquad (ii)$$

where the symbol $\tilde{a}$ denotes the matrix obtained from the matrix a by transposing rows and columns. If ψ_m, ϕ_n are suitably normalized solutions of (i) and (ii) respectively, referring to the states m and n, we take $\phi_n\psi_m$ to be the corresponding matrix element of the q's having specified values. If H is ~~Hamil~~ Hermitian, $\tilde{H}$ is the conjugate imaginary of H (obtained by writing $-i$ for i) and the solutions of (ii) are just the conjugate imaginaries to the solutions of (i), so that in this case our probability $\phi_n\psi_m$ becomes the usual one $\bar{\psi}_n\psi_m$. In the general case it is necessary to use the transposed Hamiltonian instead of the conjugate imaginary Hamiltonian in (ii) in order to secure that if ϕ_n, ψ_m are initially orthogonal or mutually normalised (i.e. $\int\phi_n\psi_m\,dq = 1$), they always remain orthogonal ~~and~~ or mutually normalised respectively.

Our wave equation for an electron in an electromagnetic field is

$$[p_0 + e'A_0 + \rho_1(\sigma, p + e'A) + \rho_3 mc]\psi = 0 \qquad (3)$$

where $e' = e/c$. The Hamiltonian here will be Hermitian if a matrix scheme for the spin variables is chosen in which they are Hermitian. However, if one now applies a Lorentz transformation to this wave equation and divides out by the coefficient of the new p_0, the resulting new Hamiltonian will not, in general, be Hermitian, although, ~~so~~ as shown in loc. cit., §3, it may be brought back to its original Hermitian form by a canonical transformation of ~~a~~ the matrix scheme for the spin variables. In the following work we require to have the same matrix representation of the spin variables for all frames of reference, so we cannot assume our Hamiltonian is Hermitian and must use the above generalized interpretation.

The equation obtained by transposing rows and columns in the ~~neighborhood of any point is given, according to the above~~ assumption, by $\phi\psi$, ~~where this product must~~ now be ~~understood~~ to mean the sum of the products of each

operator of (3) is

$$[-p_0 + e'A_0 + \tilde{\rho}_1(\tilde{\sigma}, -p + e'A) + \tilde{\rho}_3 mc]\phi = 0 \qquad (4)$$

The probability per unit volume of the electron being in the neighborhood of any point is given, according to the above assumption, by $\phi\psi$, where this product must now be understood to mean the sum of the products of each of the four components of ϕ (referring respectively to the four rows or columns of the matrices ρ, σ) into the corresponding component of ψ. We have to prove that this probability is the time component of a 4-vector, and that the divergence of this 4-vector vanishes.

From (3)

$$[\rho_3(p_0 + e'A_0) + \rho_1\rho_3(\sigma, p + e'A) + mc]\rho_3\psi = 0$$

or

$$[\gamma_0(p_0 + e'A_0) + \Sigma_{r=1,2,3}\,\gamma_r(p_r + e'A_r) + mc]\chi = 0. \qquad (5)$$

where $\gamma_0 = \rho_3$, $\quad \gamma_r = \rho_1\rho_3\sigma_r$, $\quad \chi = \rho_3\psi$.

Equation (5) is symmetrical between the four dimensions of space and time, and shows that $\gamma_0, -\gamma_1, -\gamma_2, -\gamma_3$ are the contravariant components of a four-vector. If we multiply (4) by $\tilde{\rho}_3$ on the left-hand side, we get

$$[\tilde{\gamma}_0(-p_0 + e'A_0) + \Sigma_r\,\tilde{\gamma}_r(-p_r + e'A_r) + mc]\phi = 0 \qquad (6)$$

since

$$\tilde{\gamma}_0 = \rho_3, \qquad \tilde{\gamma}_r = \tilde{\sigma}_r\tilde{\rho}_3\tilde{\rho}_1 = \tilde{\rho}_3\cancel{\rho}\tilde{\rho}_1\tilde{\sigma}_r$$

The operator is this equation is just the transposed operator of (5). The probability per unit volume of the electron being in any place is now given by

$$\phi\psi = \phi\rho_3\chi = \phi\gamma_0\chi, \qquad (7)$$

where $\phi\alpha\chi$ denotes the sum of the products of each of component of ϕ into the corresponding components of $\alpha\chi$, α being any function of the spin variables, represented by a matrix with four rows and columns. [Note that quite generally $\phi\alpha\chi = \chi\tilde{\alpha}\phi$] Expression (7) is the time component of a 4-vector, whose spacial components, namely,

$$-\phi\gamma_1\chi, \quad -\phi\gamma_2\chi, \quad -\phi\gamma_3\chi,$$

must give 1/c times the probability per unit time of the electron crossing unit area perpendicular to each of ~~the~~ three axes respectively.

We must now show that the ~~degrees~~ divergence of this 4-vector vanishes, i.e. that

$$\frac{1}{c}\frac{\partial}{\partial t}(\phi\gamma_0\chi) - \Sigma_r\frac{\partial}{\partial x_r}(\phi\gamma_r\chi) = 0 \qquad (8)$$

Multiplying (5) by ϕ and (6) by χ and subtracting, we get

$$\phi[\gamma_0 p_0 + \Sigma_r\gamma_r p_r]\chi + \chi[\tilde{\gamma}_0 p_0 + \Sigma_r\tilde{\gamma}_r p_r]\phi = 0$$

which gives

$$\phi[\gamma_0\frac{\partial}{c\partial t} - \Sigma_r\gamma_r\frac{\partial}{\partial x_r}]\chi + \chi[\tilde{\gamma}_0\frac{\partial}{c\partial t} - \Sigma_r\tilde{\gamma}_r\frac{\partial}{\partial x_r}]\phi = 0,$$

or

$$\phi[\gamma_0\frac{\partial}{c\partial t} - \Sigma_r\gamma_r\frac{\partial}{\partial x_r}]\chi + \frac{1}{c}\frac{\partial\phi}{\partial t}\gamma_0\chi - \Sigma_r\frac{\partial\phi}{\partial x_r}\gamma_r\chi = 0.$$

This gives immediately the conservation equation (8) as the γ's are here constant matrices.

§2. Selection Principle.

In loc. cit. the quantum number j was introduced, which determines the magnitude of the resultant angular momentum for an electron moving in a central field of force. j can take both positive and negative integral values. Again the magnetic quantum number $u = M_3/h$, say, that determines the component of the total angular momentum in some specified direction, was shown to take half odd integral values from $-|j| + \frac{1}{2}$ to $|j| - \frac{1}{2}$. The state $j = 0$ is thus excluded, and the weight of any state j is $2|j|$. The equation obtained to determine the energy levels, i.e., equation (25) or (26), involves j only through the combination $j(j+1)$ except in the last term, which represents the spin correction. Thus two values of j which gives the same value for $j(j+1)$ form a spin doublet, so that $j = j'$ and $j = -(j'+1)$ form

a spin doublet when $j' > 0$. The connection between j-values and the usual notation for alkali spectra is therefore given by the following scheme :—

$$j = \underset{S}{-1} \qquad \underbrace{1 \quad -2}_{P} \qquad \underbrace{2 \quad -3}_{D} \qquad \underbrace{3 \quad -4}_{D} \quad \cdots$$

There is no azimuthal quantum number k in the present theory, an orbit for an electron in an atom being defined by three quantum numbers n, j, u only. One might on this account expect the selection rules, the relative intensities of the lines of a multiplet, etc, in the usual derivation of which k plays an important part, to be different in the present theory, but it will be found that they just happen to be the same.

We shall first determine the selection rule for j. We use the following two theorems :—

(i) If a dynamical variable X anticommutes with j, its matrix elements all refer to transitions of the type $j \rightarrow -j$.

(ii) If a dynamical variable Y satisfies

$$[[Y, jh], jh] = -Y, \qquad (9)$$

its matrix elements all refer to transitions of the type $j \rightarrow j \pm 1$.

To prove (i) we observe that the condition $jX + Xj = 0$ gives

$$j'X(j'j'') + X(j'j'')\cdot j'' = 0$$

or

$$(j' + j'')X(j'j'') = 0.$$

Hence $X(j'j'') = 0$ unless $j'' = -j'$.

A proof of (ii) involving angle variables has been given in a previous paper.* A simple proof analogue to the foregoing proof of (i) is as follows. Equation (9) gives

$$Yj^2 - 2jYj + j^2Y = Y$$

or

$$Y(j'j'')j''^2 - 2j'Y(j'j'')j'' + j'^2Y(j'j'') = Y(j'j'').$$

---

* III. p. 281

Hence $\Upsilon(j'j'') = 0$ except when

$$j''^2 - 2j'j'' + j'^2 = 1,$$

i.e. when

$$j'' = j' \pm 1,$$

We shall now evaluate $[[x_3, jh], jh]$. The definition of j is

$$jh = \rho_3 \{(\boldsymbol{\sigma}, \mathbf{m}) + h\}.$$

Hence

$$[x_3, jh] = \rho_3 \{ \sigma_1 [x_3, m_1] + \sigma_2 [x_3\, m_2]\}$$

$$= \rho_3 (\sigma_1 x_2 - \sigma_2 x_1).$$

so that

$$[[x_3, jh], jh] = [\sigma_1 x_2 - \sigma_2 x_1, (\boldsymbol{\sigma}\, \mathbf{m})]$$

Now

$$ih[\sigma_1, (\boldsymbol{\sigma}, \mathbf{m})] = \sigma_1 (\boldsymbol{\sigma}, \mathbf{m}) - (\boldsymbol{\sigma}, \mathbf{m})\sigma_1$$

$$= 2i(\sigma_3 m_2 - \sigma_2 m_3)$$

or

$$\tfrac{1}{2} h [\sigma_1, (\boldsymbol{\sigma}\, \mathbf{m})] = \sigma_3 m_2 - \sigma_2 m_3$$

and similarly

$$\tfrac{1}{2} h [\sigma_2, (\boldsymbol{\sigma}, \mathbf{m})] = \sigma_1 m_3 - \sigma_3 m_1 .$$

Hence

$$\tfrac{1}{2} h [[x_3, jh], jh] = (\sigma_3 m_2 - \sigma_2 m_3) x_2 + \tfrac{1}{2} h \sigma_1 (\sigma_3 x_1 - \sigma_1 x_3)$$

$$-(\sigma_1 m_3 - \sigma_3 m_1) x_1 - \tfrac{1}{2} h \sigma_2 (\sigma_2 x_3 - \sigma_3 x_2)$$

$$= \sigma_3 (\mathbf{m}, \mathbf{x}) - m_3 (\boldsymbol{\sigma}, \mathbf{x}) + \tfrac{1}{2} h \{-\sigma_3 (\boldsymbol{\sigma}, \mathbf{x}) - x_3\}$$

$$= - M_3 (\boldsymbol{\sigma}, \mathbf{x}) - \tfrac{1}{2} h x_3 ,$$

so that

$$[[x_3, jh], jh] = -2u(\boldsymbol{\sigma}, \mathbf{x}) - x_3 .$$

Thus x_3 does not quite satisfy the conditions that Υ satisfies in (9), owing the extra term $-2u(\boldsymbol{\sigma}, \mathbf{x})$. This extra term, however, anticommutes with j. If we now form the expression $x_3 - cu(\boldsymbol{\sigma}, \mathbf{x})$ where c is some quantity that commutes with j, we can choose c so as to make this expression satisfy completely the condition that Υ satisfies in (9). We have, in fact,

$$[[x_3 - cu(\boldsymbol{\sigma}, \mathbf{x}), jh], jh] = -2u(\boldsymbol{\sigma}, \mathbf{x}) - x_3 + cu4j^2(\boldsymbol{\sigma}, \mathbf{x})$$

$$= -\{x_3 - cu(\boldsymbol{\sigma}, \mathbf{x})\}$$

if c is chosen such that

$$-2 + 4j^2 c = c ,$$

i.e. if

$$c = \tfrac{1}{2}(j^2 - \tfrac{1}{4}).$$

Hence x_3 can be expressed as the sum of two ~~electrons~~ terms, namely,

$$\frac{u}{2(j^2 - \frac{1}{4})}(\boldsymbol{\sigma}, \mathbf{x}) \quad \text{and} \quad x_3 - \frac{u}{2(j^2 - \frac{1}{4})}(\boldsymbol{\sigma}, \mathbf{x}),$$

of which the first anticommutes with j, and therefore contains only matrix elements referring to transitions of the type $j \to -j$, while the second satisfies the condition that Υ satisfies in (9), and therefore contains only matrix elements referring to transitions of the types $j \to j \pm 1$. A similar result holds for x_1 and x_2. Hence the selection rule for j is

$$j \to -j \quad \text{or} \quad j \to j \pm 1 .$$

Thus from states with $j = 2$ transitions can take place to states with $j = 1$, -2 or 3. Comparing this selection rule with the above scheme connecting j-values with the S, P, D notation, we see that it is exact equivalent to the two selection rules for j and k of the usual theory, and is therefore in agreement with experiment.

§ The Relative Intensities of the Lines of a Multiplet

The relative intensities of the various components into which a line is split up in a weak magnetic field must be the same on the present theory as on previous theories, as they depend only on the Vertauschungs relations connecting the co-ordinates x_r with the components of total angular momentum M_r, which are taken over unchanged into the present theory. It will therefore be sufficient, for determining the relative intensities of the lines of a multiplet, to consider only one Zeeman component of each line, say, the component for which $\Delta u = 0$, i.e. the component that comes from x_3.

We shall determine the matrix elements of x_3, when expressed as a matrix in a scheme in which r, j, u and ρ_3 are diagonal. x_3 is diagonal in (i.e., commutes with) all of these variables except j. The part of x_3 referring to transitions $j \to -j$ we found to be

$$\frac{u}{2(j^2-\frac{1}{4})}(\sigma, x) = \frac{u}{2(j^2-\frac{1}{4})}\varepsilon\rho_1 r. \qquad (11)$$

using the ε introduced in loc. cit. §6. $\varepsilon\rho_1$ anticommutes with j, so that it can contain only matrix elements of the type $\varepsilon\rho_1(j,-j)$, and from the condition $(\varepsilon\rho_1)^2 = 1$ we must have

$$|\varepsilon \cancel{p}\rho_1, (j,-j)| = 1.$$

Hence $$|x_3, (j,-j)| = \frac{u}{2(j^2-\frac{1}{4})} r\, |\varepsilon\rho_1 (j,-j)| = \frac{u}{2(j^2-\frac{1}{4})} r. \qquad (12)$$

Again, we have from (10)

$$\{x_3 - i[x_3, jh]\}\{x_3 + i[x_3, jh]\} = \{x_3 - i\rho_3(\sigma_1 x_2 - \sigma_2 x_1)\}$$
$$\times\{x_3 + i\rho_3(\sigma_1 x_2 - \sigma_2 x_1)\}$$
$$= x_3^2 + (\sigma_1 x_2 - \sigma_2 x_1)^2 = r^2,$$

which gives

$$\{(j+1)x_3 - x_3 j\}\{x_3(j+1) - j x_3\} = r^2$$

If we equate the (j, j) matrix elements of each side of this equation, we get on the left-hand side the sum of three terms, namely, the $(j,-j)$ ~~$(j, j+1)$~~ matrix element of the 1st $\{\ \}$ bracket times the $(-j, j)$ ~~$(j+1, j)$~~ element of the second, ~~and~~ the $(j, j+1)$ element of the first times the $(j+1, j)$ element of the second, and the $(j, j-1)$ element of the first times the $(j-1, j)$ element of the second. The sound of these three terms vanishes, leaving

$$(2j+1)^2|x_3(j,-j)|^2 + 4|x_3(j, j-1)|^2 = r^2.$$

Hence

$$|x_3(j, j-1)|^2 = \frac{1}{4}r^2\left\{1 - \frac{u^2}{(j-\frac{1}{2})^2}\right\} = \frac{1}{4}r^2\frac{(j+u-\frac{1}{2})(j-u-\frac{1}{2})}{(j-\frac{1}{2})^2} \qquad (13)$$

Writing $-j$ for j, we get

$$|x_3(-j, -j-1)|^2 = \frac{1}{4}r^2\frac{(j+u+\frac{1}{2})(j-u+\frac{1}{2})}{(j+\frac{1}{2})^2}. \qquad (14)$$

The three matrix elements of x_3 given in (12), (13) and (14) are associated with the three components of the multiplet formed by the combination

of two doublets. The ratio of these matrix elements will, to a first approximation, remain unchanged when one makes a transformation from the matrix scheme in which r, j, u, ρ_3 are diagonal to a scheme in which the Hamiltonian is diagonal, and will therefore gives the relative intensities of the Zeeman components $\Delta u = 0$ of the lines in a combination doublet. These radios are in agreement with those of ~~the~~ previous theories based on the spinning electron model.

§4. The Zeeman Effect

If there is a uniform magnetic field of intensity H in the direction of the x_3 axis, we can take the magnetic potentials to be

$$A_1 = -\tfrac{1}{2}Hx_2, \qquad A_2 = \tfrac{1}{2}Hx_1, \qquad A_3 = 0$$

The additional terms appearing in the Hamiltonian F will now be

$$\Delta F = \rho_1 e'(\sigma, A) = -\tfrac{1}{2}He'\rho_1(\sigma_1 x_2 - \sigma_2 x_1).$$

From (10) it follows that $\rho_3(\sigma_1 x_2 - \sigma_2 x_1)$ or $(\sigma_1 x_2 - \sigma_2 x_1)$, like x_3, contains only matrix elements of the types $(j,-j)$ or $(j, j\pm1)$. Now ρ_1 anticommutes with j, and therefore contains only matrix elements of the type $(j,-j)$. Hence ΔF contains only matrix elements of the type (j, j) or $(j, -j\pm1)$.

In loc. cit. §6, it was found [see equation (24)] that the Hamiltonian could be expressed as

$$F \equiv p_0 + V + \varepsilon p_r + i\varepsilon\rho_3\, jh/r + \rho_3 mc. \qquad (15)$$

It follows from (10) that $\rho_2(\sigma_1 x_2 - \sigma_2 x_1)$ anticommutes with (σ, x) and therefore also with ε. Thus if we put

$$\Delta F = ih\varepsilon\rho_3\eta r.$$

so that $$\eta = \tfrac{1}{2}He/ch \cdot \varepsilon\rho_2(\sigma_1 x_2 - \sigma_2 x_1)/r,$$

η commutes with ε. Further, η commutes with ρ_3, r and p_r, so that it commutes with all the variables occurring in (15) except j. If we now express η as a matrix in j, we shall have obtained an expression for $\Delta \bar{F}$ in terms of the variables occurring in (15). We have, from (10) and (13)

$$|\rho_3(\sigma_1 x_2 - \sigma_2 x_1)\,(j, j \underset{-}{\cancel{+}} 1)|^2 = |x_3(j, j\underset{-}{\cancel{+}}1)|^2 = \frac{1}{4} r^2 \frac{(j+u-\frac{1}{2})(j-u-\frac{1}{2})}{(j-\frac{1}{2})^2},$$

and similarly

$$|\rho_3(\sigma_1 x_2 - \sigma_2 x_1)\,(j, j+1)|^2 = |x_3(j, j+1)|^2 = \frac{1}{4} r^2 \frac{(j+u+\frac{1}{2})(j-u+\frac{1}{2})}{(j+\frac{1}{2})^2}$$

We have seen that the matrix elements of $\varepsilon\rho_1$, all of which are of the type $(j, -j)$, must be of modulus unity. Hence

$$|\eta(j, -j-1)|^2 = \left(\frac{He}{2ch r}\right)^2 |i\varepsilon\rho_1(j,-j)|^2 |\rho_3(\sigma_1 x_2 - \cancel{x_2}\,\sigma_2 x_1)(-j, -j-1)|^2$$

$$= \left(\frac{He}{4ch}\right)^2 \frac{(j+u+\frac{1}{2})(j-u+\frac{1}{2})}{(j+\frac{1}{2})^2} \qquad \Bigg\} (16)$$

and similarly

$$|\eta(j, -j+1)|^2 = \left(\frac{He}{4ch}\right)^2 \frac{(j+u-\frac{1}{2})(j-u-\frac{1}{2})}{(j-\frac{1}{2})^2}$$

Again, from (10) and (11)

$$\rho_3(\sigma_1 x_2 - \sigma_2 x_1)(-j, j) = -2ij \cdot x_3(-j, j) = \frac{-u}{(j^2-\frac{1}{4})}\, irj\,(\varepsilon\rho_1)(-j, j),$$

so that

$$\eta(j, j) = \frac{He}{2ch} \frac{uj}{j^2-\frac{1}{4}} \qquad (17)$$

If we write down in full, as in loc. cit., the wave equation corresponding to (15), and include the extra term ΔF, we shall have

$$[(F+\Delta F)\psi]_\alpha = (p_0+V)\psi_\alpha - h\frac{\partial}{\partial r}\psi_\beta - \left(\frac{j}{r}+\eta r\right)h\psi_\beta + mc\psi_\alpha = 0,$$

$$[(F+\Delta F)\cdot\psi]_\beta = (p_0+V)\psi_\beta + h\frac{\partial}{\partial r}\psi_\alpha - \left(\frac{j}{r}+\eta r\right)h\psi_\alpha - mc\psi_\beta = 0$$

where η is now an operator, operating on ψ_α and ψ_β, that commutes with everything except j. On eliminating ψ_α, this gives, corresponding to (25) of loc. cit.,

$$\frac{\partial^2}{\partial r^2}\psi_\beta + \left[\frac{(p_0+V)^2 - m^2c^2}{h^2} - \frac{j(j+1)}{r^2} + \eta - \eta j - j\eta - \eta^2 r^2\right]\psi_\beta$$

$$-\frac{1}{p_0+V+mc}\frac{\partial V}{\partial r}\left[\frac{\partial}{\partial r} + \frac{j}{r} + \eta r\right]\psi_\beta = 0$$

We can neglect the $\eta^2 r^2$ term, which is proportional to the square of the field strength, and also the ηr term in the last bracket, which is of the order of magnitude of field strength times spinning correction. The only first order effect of the field is the insertion of the terms $\eta - \eta j - j\eta$ in the 1st bracket. This bracket may ~~be~~ now be written as

$$\left[\frac{2mE}{h^2} + \frac{E^2}{c^2h^2} + \frac{2(E+mc^2)}{ch^2}V + \frac{V^2}{h} - \frac{j(j+1)}{r^2} + \eta - \eta j - j\eta\right], \qquad (18)$$

where E is the energy level, equal to $p_0c - mc^2$.

If the field is weak compared with the doublet separation, we can obtain a first approximation to the change in the energy levels by neglecting the non diagonal matrix elements of ΔF or of η. The extra terms $\eta - \eta j - j\eta$ in (18) are now a constant instead of an operator, namely, the constant

$$-(2j-1)\,\eta(j, j) = -\frac{He}{ch}\frac{uj}{j+\frac{1}{2}}$$

from (17). The energy levels will be reduced by $h^2/2m$ times this constant, if we neglect the fact that the characteristic E occurs in (18) in other places besides the terms $2mE/h^2$, which means neglecting the interaction of the magnetic field with the relativity variation of mass with velocity. The increase in the energy levels caused by the magnetic field is thus

$$\frac{He}{2mc}\frac{j}{j+\frac{1}{2}}\,uh = \omega g u h$$

where ω is the Larmor frequency $He/2mc$, and g, the Landé splitting factor, has the value

$$g = j/(j+\tfrac{1}{2}),$$

For the succession of j-values, $-1, 1, -2, 2, -3 \ldots$ g has the values, $2, \frac{2}{3}, \frac{4}{3}, \frac{4}{5}, \frac{6}{5}$; in agreement with Landé's formula for alkali spectra.

We now take the case of a magnetic field that is strong compared with the doublet separation, but weak compared with the ~~doublet system~~ separation of terms of different series. This requires that the matrix elements of η of the type $\eta(j,-j-1)$ with $j>0$ shall be taken into account, although those of the type $\eta(j,-j+1)$ can still be neglected. The reduction in the energy levels will now be approximately $h^2/2m$ times one or other of the characteristic values of the extra terms

$$\eta-\eta j-j\eta$$

in (18). These characteristic values are the roots ζ of the equation

$$\begin{vmatrix} (\eta-\eta j-j\eta)(j,j)-\zeta & (\eta-\eta j-j\eta)(j,-j-1) \\ (\eta-\eta j-j\eta)(-j-1,j) & (\eta-\eta j-j\eta)(-j-1,-j-1)-\zeta \end{vmatrix}=0.$$

or

$$\begin{vmatrix} -(2j-1)\cdot\eta(j,j)-\zeta & 2\eta(j,-j-1) \\ 2\eta(-j-1,j) & (2j+3)\cdot\eta(-j-1,-j-1)-\zeta \end{vmatrix}=0.$$

This gives, with the help of (16) and (17)

$$\zeta^2+\frac{He}{ch}\left[\frac{uj}{j+\frac{1}{2}}+\frac{u(j+1)}{j+\frac{1}{2}}\right]\zeta+\left(\frac{He}{ch}\right)^2\left[\frac{u^2j(j+1)}{(j+\frac{1}{2})^2}-\frac{(j+\frac{1}{2})^2-u^2}{4(j+\frac{1}{2})^2}\right]=0,$$

which reduces to

$$\zeta^2+\frac{He}{ch}2u\zeta+\left(\frac{He}{ch}\right)^2\left(u^2-\frac{1}{4}\right)=0$$

Hence
$$\zeta=-\frac{He}{ch}\left(u\pm\frac{1}{2}\right).$$

The increase in the energy levels due to the magnetic field is therefore

$$-\frac{h^2}{2m}\zeta=\frac{h^2}{2m}\frac{He}{m}\left(u\pm\frac{1}{2}\right)=\omega\left(u\pm\frac{1}{2}\right)h,$$

in agreement with the previous spinning electron theory of the Paschen-Back effect.

One might expect that with still stronger magnetic fields the matrix elements $(j,-j+1)$ of η would come into play, and would cause interference between the Zeeman pattern ~~and~~ of terms whose quantum numbers k in the usual notation differ by 2. The matrix elements $(j,-j+1)$ of $\eta-\eta j-j\eta$, however, vanish for arbitrary η, so that no effect of this nature occurs.

Zum Mehrkörperproblem der Quantentheorie

P. Jordan und O. Klein. (<u>45</u> S. 751)

Die sogenannte Wellengleichung im Koordinatenraum (mit den entsprechenden matrizenmechanischen Bewegungsgleichungen) bildet bekanntlich eine Umschreibung des Mehrkörperproblems der klassischen Punktmechanik, die allen Forderung der Quantentheorie genügt, und die zu einer mit der Erfahrung übereinstimmenden quantitativen Behandlung der wichtigsten Probleme des Atombaues geführt hat. Es liegt jedoch in der Natur der Sache, daß das Anwendungsgebiet dieser Methode begrenzt ist, indem die klassischen Bewegungsgleichungen einer Anzahl elektrischer Teilchen nur dann auf Hamiltonschen Form gebracht ~~ist~~ werden können, wenn die endliche Ausbreitung der elektromagnetischen Felder vernachlässigt werden kann.

Es ist deshalb noch nicht möglich gewesen, eine dem Wesen der Relativitätstheorie entsprechende Behandlung des Quantenmech. Mehrkörperproblems zu finden. In der klassischen Theorie wird die in Frage stehende Schwierlichkeit bekanntlich dadurch überwunden, daß neben den Koordinaten der Teilchen die sogenannten Feldgröße eingeführt werden, die besonderen Gleichungen, den Feldgleichungen, zu genügen haben, welche mit den Bewegungsgleichungen zusammen einem verallgemeinerten Hamiltonschen Prinzip gehorchen. In Analogie hierzu wird man wohl eine rationelle Lösung des allgemeinen Mehrkörperproblems der Quantentheorie erst in Zusammenhang mit einer Revision der gesamten Feldtheorie im Sinne der Quantenpostulate erwarten dürfen.

Grundlegend für die Bestrebung in dieser Richtung war bekanntlich die von Einstein begründete Lichtquantentheorie, durch welche der in diesem Zusammenhang besonders wichtige Dualismus „Korpuskel-Welle" zuerst in die Physik eingeführt wurde. Führ das Mehrkörperproblem kommen hierbei Einsteins Arbeiten über Gasentartung in Betracht, wo er im Anschluß an de Broglie ein ideales Gas mit einem System von gequantelten Wellen im dreidimensionalen Raume vergleichen hat[1]. Dieser Gesichtspunkt dürfte deshalb geeignet sein zu einem Angriff auf das relativistische Mehrkörperproblem, weil hierbei die Materie und das Feld mathematischen in gleichartiger Weise beschrieben werden, nämlich durch partielle Differentialgleichungen. In der Tat besitzen wir in der Wellenmechanik des Einkörperproblems in seinem Zusammenhang mit der Feldtheorie ein „klassisches Modell", das ~~eben~~ den Forderungen der Relativitätstheorie genügt, und es liegt deshalb nahe, zu versuchen, ob dieses als Grundlage für das Mehrkörperproblem benutzt werden kann, indem man die darin auftretenden Grössen — elektromagnetische Potentiale und Schrödingerschen Wellenfunktionen — einer „Quantelung" unterwirft. Als Vorbereitung zu einem solchen Versuch zeigen wir in der vorliegenden Arbeit, dass es möglich ist, von diesem Gesichtspunkt aus in dem Spezialfalle, wo die endliche Ausbreitungsgeschwindigkeit der Felder vernachlässigt wird, zu Resultaten zu gelangen, die mit der Wellengleichung im Koordinatenraum übereinstimmen. Wir schliessen uns hierbei einer Arbeit von ~~energetische unabhängigen Teilchen~~ gelöst wurde[2]. Dirac an, wo das entsprechende Problem für eine Anzahl von energetisch unabhängigen Teilchen gelöst wurde[2]. Wie Dirac, beschränken wir uns auf den Fall, wo die Eigenfunktionen im Koordinatenraum symmetrisch in den Teilchen sind, obgleich die Ergebnisse im Hinblick auf die von ~~Dirac Betrachtungen wohl auch im~~ dem einen der Verfasser 3) neulich gegebene Verallgemeinerung von Diracs Betrachtungen wohl auch im antisymmetrischen Falle Gültigkeit haben dürfen.

Mit dem Problem einer einheitlichen Behandlung von Licht und Materiewellen hängen übrigens offensichtlich auch gewisse grundsätzliche Schwierigkeiten zusammen, auf die Schrödinger[4] mit besonderem Nachdruck hingewiesen hat, und die darin zutage treten, dass es mit den bislang gebrauchte Hilfsmitteln — ohne eine „Quantelung" der Wellen — nicht möglich ist, ein „geschlossenes" System von Feldgleichungen aufzustellen, welches die Schrödingersche Theorie des Wasserstoffatoms in sich enthält. Insbesondere einige Tatsachen, zu deren Besprechung wir in §3 dieser Arbeit kommen, scheinen darauf hinzudeuten, dass auch für von Schrödinger erörterten Schwier~~tich~~igkeiten eine naturgemässe Lösung zu finden sein wird durch die Fortführung derjenigen Theorie, zu der wir hier einen Beitrag liefern.

§1. Quantelung der de Brogliewellen im Falle elektrischer Wechselwirkung. Wir wollen ein System von Elektronenwellen betrachten, die teils von einem äusseren elektrostatischen Felde ~~zu~~ und teils von dem nach der Schrödingerschen Dichtehypothese berechneten Eigenfeld beeinflusst werden. Hierbei wollen wir sowohl die magnetischen Wechselwirkungen als auch den Einfluss der endlichen Fortpflanzungsgeschwindigkeit des elektrischen Feldes und die Relativitätskorrektion vernachlässigten. Wir werden später diese Annahmen wesentlich verallgemeinern und insbesondere auch magnetische Kräfte mit berücksichtigen können. Um Konvergenzschwierigkeiten zu vermeiden, wollen wir ferner annehmen, dass die Wellen sich in einem endlichen geschlossenen Gefäss befinden, an dessen Wänden die Wellenamplituden und das elektrische Potential verschwinden mögen. Einen Punkt im Raume bezeichnen wir durch den von irgend einem festen Punkt zu dem gezogenen Radiusvektor $\mathfrak{r}$. Eine Funktion φ des Punktes $\mathfrak{r}$ im Raume bezeichnen wir dementsprechend mit $\varphi(\mathfrak{r})$. Es seien nun

---

1. Vgl. hierzu Born, Heisenberg und Jordan über die Schwankungseigenschaften der Strahlung, ferner und Dirac., sowie Pauli, Jordan: Invariante Quantenelektrodynamik.
2. Dirac 114. 243
3. Jordan 44. 473, (Die Überlegungen dieser Arbeit sind in zwei Punkten nicht korrekt.)

4. 82 265
1. Wir beschränken uns hier auf den Fall nur gleichartiger Teilchen; es bieten jedoch keine Schwierigkeit, in ähnlicher Weise mehrere Arten von Teilchen, etwa Protonen*

* und Elektronen, zu berücksichtigen.

$\varphi(r,t)$ und $\varphi^*(r,t)$ die konjugiert komplexen Wellenamplituden, als Funktionen von r und der Zeit t betrachtet, und $V(r,t)$ bezeichne das elektrische Potential. Die Funktionen φ und φ^* sollen dann, abgesehen von einzelnen Punkten, wo das Potential des äußeren Feldes singulär wird (Punktladungen), folgende Differentialgleichungen befriedigen:

$$\left.\begin{aligned} -\frac{h^2}{8\pi^2\mu}\Delta\varphi + \frac{h}{2\pi i}\frac{\partial\varphi}{\partial t} - \varepsilon V\varphi = 0, \\ -\frac{h^2}{8\pi^2\mu}\Delta\varphi^* - \frac{h}{2\pi i}\frac{\partial\varphi^*}{\partial t} - \varepsilon V\varphi^* = 0, \\ \Delta V = 4\pi\varepsilon\,\varphi^*\varphi \end{aligned}\right\} \quad (1)$$

Hier bedeutet h die Plancksche Konstante, μ die Masse und und $-\varepsilon$ die Ladung des Elektrons, während Δ den Laplaceschen Differentialoperator bezeichnet[1]. Nach der dritten Gleichung können wir nun setzen:

$$V(r) = V_0(r) - \varepsilon\int G(r,r')\,\varphi^*(r')\,\varphi(r')\,dv', \qquad (2)$$

wo ~~v_0~~ $V_0(r)$ das Potential des „äußeren" Feldes im Punkte r bezeichnet, während $G(r,r')$ die Greensche Funktion der Laplaceschen Gleichung des von dem Gefäß bezeichneten Raumgebietes vorstellen soll und dv' ein Raumdifferential in der Umgebung des Punktes r' bedeutet. Die Integration ist über dieses ganze Gebiet zu erstrecken. Wenn das Gefäß groß ist, gilt bekanntlich mit großer Annäherung

$$G(r,r') = \frac{1}{|r-r'|}, \qquad (3)$$

wo $|r-r'|$ die Länge des Vektors $r-r'$ bezeichnet, d. h. die Entfernung der Punkte r und r'.

Das Gleichungssystem (2) läßt sich bekanntlich aus einem Hamiltonschen Prinzip ableiten, dem wir die folgende Form geben können:

$$\delta\int_{t_0}^{t} dt\int L\,dv = c \qquad (4)$$

wo

$$L = \frac{h^2}{8\pi^2\mu}(\operatorname{grad}\varphi^*\cdot\operatorname{grad}\varphi) + \frac{h}{2\pi i}\left(\varphi^*\frac{\partial\varphi}{\partial t} - \frac{\partial\varphi^*}{\partial\varphi}\varphi\right) - \varepsilon V\varphi^*\varphi - \frac{1}{8\pi}(\operatorname{grad}(V-V_0))^2, \qquad (5)$$

$-\varepsilon\varphi^*V\varphi$

$\frac{h^2}{8\pi^2\mu}$ grad φ^* (div grad $\varphi\cdot\delta s$ + grad $\frac{\partial\varphi}{\partial t}$ δ( ) + grad $\frac{\partial\varphi}{\partial t}\delta t$)

$+\varphi^*$(grad $\frac{\partial\varphi}{\partial t}\cdot\delta s$ + gra $\frac{\partial^2\varphi}{\partial t^2}\delta t$) +

$-\varepsilon V\varphi^*$(grad

und wo die Variationen von φ^*, φ und am Rande des fest gedachten raumzeitlichen Integrationsgebietes verschwinden sollen.

Es sollen $u_s(r)$ die normierten Eigenfunktionen der Gleichung

$$\frac{h^2}{8\pi^2\mu}\Delta u + (E+\varepsilon V_0)\,u = 0 \qquad (6)$$

und E_s seien die ~~F~~entsprechenden Eigenwerte, d. h. die Energiewerte des Systems, wenn das äußere Feld allein vorhanden ist. Wir können nun die Funktionen φ und φ^* nach den Eigenfunktionen u_s entwickeln, indem wir setzen

$$\left.\begin{aligned} \varphi(r,t) = \sum_s b_s(t)\,u_s(r) \\ \varphi^*(r,t) = \sum_s b_s^*(t)\,u_s(r) \end{aligned}\right\} \quad \ldots\ldots (7)$$

wo b_s und b_s^* zueinander konjugiert komplexe Funktionen der Zeit allein sind. Führen wir diese Ausdrücke in L ein, so ergibt sich unter Berücksichtigung von (2) und (6) auf Grund der Normierung der Eigenfunktionen $u_s(r)$:

$$M = \int L\,dv = \sum_s b_s^* b_s E_s + \frac{\varepsilon^2}{2}\sum_{\mu\nu rs} A(\mu\nu|rs)\,b_r^* b_s^* b_\mu b_\nu + \frac{h}{4\pi i}\sum_s\left(b_s^*\frac{db_s}{dt} - \frac{db_s^*}{dt}b_s\right), \quad \ldots\ldots (8)$$

wo

$$A(\mu\nu|rs) = \iint G(r',r'')\,u_\mu(r')\,u_\nu(r'')\,u_r(r')\,u_s(r'')\,dv'\,dv'' \qquad (9)$$

Das Variationsprinzip (4) nimmt nun die Form

$$\delta\int M\,dt = 0 \qquad (10)$$

an, welches die Form des Hamiltonschen Prinzips der Mechanik ist.

Berechnen wir demnächst die Totalenergie des Systems. Nach bekannten Sätzen der Wellenmechanik können wir setzen

$$E = \int dv\left\{\frac{h^2}{8\pi^2\mu}(\operatorname{grad}\varphi^*\cdot\operatorname{grad}\varphi) - \varepsilon V_0\varphi^*\varphi + \frac{1}{8\pi}(\operatorname{grad}(V-V_0))^2\right\}, \qquad (11)$$

woraus folgt:

$$E = \sum_s b_s^* b_s E_s + \frac{\varepsilon^2}{2}\sum_{\mu\nu rs} A(\mu\nu|rs)\,b_r^* b_s^* b_\mu b_\nu. \qquad (12)$$

$$\Sigma b^\dagger_\mu \varphi_\mu \Sigma b_\mu \varphi_\mu \qquad b^\dagger_\mu b_\mu = N$$

Durch Vergleich mit (8) folgen nun aus (10) die folgenden kanonischen Gleichungen für b_s und b_s^*:

$$\frac{h}{2\pi i}\frac{db_s}{dt} = -\frac{\partial E}{\partial b_s^*} \qquad \frac{h}{2\pi i}\frac{db_s^*}{dt} = \frac{\partial E}{\partial b_s} \qquad (13)$$

wonach $\frac{h}{2\pi i} b_s$ und b_s^* ein Paar kanonische Veränderliche darstellen.

Wegen des letzten Resultats können wir nun nach dem Vorgang von Dirac die Größen $\frac{h}{2\pi i} b_s$ und b_s^* durch konjugierte q-Zahlen $\frac{h}{2\pi i} b_s$, $b_s^\dagger$ ersetzen[1]. Die Größen b_s und $b_s^\dagger$ sollen demnach nicht „vertauschbar" sein, sondern der Relation

$$b_\nu^\dagger \frac{h}{2\pi i} b_\mu - \frac{h}{2\pi i} b_\mu \cdot b_\nu^\dagger = \frac{ih}{2\pi}\delta_{\mu\nu}$$

oder

$$b_\mu b_\nu^\dagger \underset{+}{-} b_\nu^\dagger b_\mu = \delta_{\mu\nu} \qquad (14)$$

unterworfen sein, wodurch die Gleichungen (13) in die folgenden q-Zahlgleichungen übergehen:

$$\frac{h}{2\pi i}\frac{db_s}{dt} = E b_s - b_s E, \qquad \frac{h}{2\pi i}\frac{db_s^\dagger}{dt} = E b_s^\dagger - b^\dagger E \qquad (15)$$

Um das Problem zu einem bestimmten zu machen, müssen wir aber noch eine Festsetzung über die Reihenfolge der Faktoren in dem Ausdruck für E treffen. Wir kommen auf diese Frage weiter unten zurück und begnügen uns für den Augenblick damit, die in (12) gewählte Reihenfolge den nun folgenden Überlegungen zugrunde zu legen.

Durch die Beziehung (14) und (15) werden bekanntlich die Forderung der Quantentheorie erfüllt; wie wir nun sehen werden, führen diese Ansätze zu der in den Teilchen symmetrischen Lösung der entsprechenden Wellengleichung im Koordinatenraum. Wie der eine der Verfasser[1] gezeigt hat, stellen die Relationen (14) jedoch nicht die einzig mögliche Lösung des Problems der Einführung der Quantentheorie im vorliegenden Fall dar, sondern es gibt eine weitere Möglichkeit, welche – jedenfalls wenn die Wechselwirkung der Teilchen vernachlässigt wird – zu dem antisymmetrischen Falle der Koordinatenraumgleichung führt.

$$\frac{\partial}{\partial x}\varphi e^\psi - \varphi\frac{\partial e^\psi}{\partial x} = e^\psi\frac{\partial\varphi}{\partial x} \qquad \frac{\partial}{\partial x}\left(e^{-ax}\cdot \psi\frac{\partial e^{ax}}{\partial x}\right) = e^{ax}\frac{\partial\psi}{\partial x}$$

Wir führen nun, Dirac[2] folgend, anstatt der komplexen Amplituden b_s und $b_s^\dagger$ die reellen Größen N_s und Θ_s ein, indem wir setzen

$$b_s = e^{-\frac{2\pi i}{h}\Theta_s} N_s^{\frac12}, \qquad b_s^\dagger = N_s^{\frac12}\cdot e^{\frac{2\pi i}{h}\Theta_s}, \qquad (16)$$

wo die Eigenwerte N_s' der Größe N_s die anschauliche Bedeutung der möglichen Anzahl von Elektronen im Zustand s bei Loslösung der gegenseitigen Kopplung der Elektronen haben[3]. Wir bekommen dann

$$(17)\qquad E = \sum_s N_s E_s + \frac{\varepsilon^2}{2}\sum_{\mu\nu r s} A(\mu\nu|rs)\, N_r^{\frac12} e^{\frac{2\pi i}{h}\Theta_r} N_s^{\frac12} e^{\frac{2\pi i}{h}(\Theta_s-\Theta_\mu)} N_\mu^{\frac12} e^{-\frac{2\pi i}{h}\Theta_\mu} N_\nu^{\frac12},$$

während die Relationen, welche die Nichtvertauschbarkeit von N_μ und $e^{\frac{2\pi i}{h}\Theta_\mu}$ ausdrücken, folgende Form annehmen:

$$N_\mu e^{\frac{2\pi i}{h}\Theta_\mu} - e^{\frac{2\pi i}{h}\Theta_\mu} N_\mu = e^{\frac{2\pi i}{h}\Theta_\mu} \qquad (18)$$

Diese Relation zeigt, daß N_μ und Θ_μ kanonisch konjugiert sind, und daß $e^{\frac{2\pi i}{h}\Theta_\mu}$ als ein Operator betrachtet werden kann, der, auf eine Funktion der Größe N' angewandt, N_μ durch $N_\mu - 1$ ersetzt.

Es sei nun $\varphi(N_1', N_2', \dots)$ diejenige Funktion der Größen N_1', $N_2' \dots$, die nach der allgemeinen Theorie von Dirac und einem der Verfasser die Bedeutung hat, daß $|\varphi|^2$ die Wahrscheinlichkeitsfunktion für die Größen N' be Loslösung der gegenseitigen Kopplung der Elektronen darstellt, und die der Hamilton-Jacobischen Funktion der klassischen Mechanik entspricht. Sie befriedigt die folgende Gleichung vom Schrödingerschen Typus:

$$-\frac{h}{2\pi i}\frac{\partial\varphi(N_1', N_2', \dots)}{\partial t} = E\varphi(N_1', N_2', \dots) \qquad (19)$$

wo E durch den Ausdruck (17) gegeben sein soll. Auf Grund der Operatoreigenschaft der Größen $e^{\frac{2\pi i}{h}\Theta_\mu}$ kann dieselbe folgendermaßen umgeformt werden. Man hat

$$N_r^{\frac12} e^{\frac{2\pi i}{h}\Theta_r} N_s^{\frac12} e^{\frac{2\pi i}{h}(\Theta_s-\Theta_\mu)} N_\mu^{\frac12} e^{-\frac{2\pi i}{h}\Theta_\nu} N_\nu^{\frac12}$$

---

[1] Über den Grund für den Gebrauch des Zeichens $b^\dagger$ statt b^* beim Übergang zu q-Zahlen siehe die Fußnote S. 756.

[1] P. Jordan, l.c.

~~2. M. Dirac, l.c.~~

[2] Dirac, l.c.

[3] Die Größe $b_s^\dagger$ geht aus b_s hervor, indem nicht nur i durch −i ersetzt, sondern außerdem die Reihenfolge der Multiplikationen umgekehrt wird; wir gebrauchen deshalb die Bezeichnung $b^\dagger$ entsprechend einem kürzlich gemachten Vorschlag, solche „Adjungierte" Größen durch ein † zu kennzeichnen.

$$= N_r^{1/2}(N_s-\delta_{rs})^{1/2}(N_\mu+1-\delta_{\mu s}-\delta_{\mu r})^{1/2}(N_\nu+1+\delta_{\mu\nu}-\delta_{s\nu}-\delta_{r\nu}) \times e^{\frac{2\pi i}{h}(\Theta_r+\Theta_r-\Theta_\mu-\Theta_\nu)} \qquad (20)$$

und also

$$-\frac{h}{2\pi i}\frac{\partial\varphi(N_1',N_2',\dots)}{\partial t} = \sum_s N_s' E_s\,\varphi(N_1',N_2',\dots)$$
$$+\frac{\varepsilon^2}{2}\sum_{\mu\nu rs} A(\mu\nu|rs)\,N_s'^{1/2}(N_r'-\delta_{rs})^{1/2}(N_\nu'+1-\delta_{\nu s}-\delta_{\nu r})^{1/2}(N_\mu'+1+\delta_{\mu\nu}-\delta_{\mu s}-\delta_{\mu r})^{1/2}e^{\frac{2\pi i}{h}(\Theta_r+\Theta_s-\Theta_\mu-\Theta_\nu)}\cdot\varphi(N_1',N_2'\dots) \qquad (21)$$

Aus (21) ersieht man, daß die Eigenwerte der Größen N_r wirklich ganze Zahlen sind. Ferner folgert man leicht aus den Gleichungen (15), daß die Größe $N=\sum_s N_s = \sum_s b_s^\dagger b_s$ zeitlich konstant ist, d.h. eine Diagonalmatrix bildet, was der Erhaltung der Elektrizität auf Grund der „klassischen" Gleichungen (1) entspricht.

Betrachten wir vorweg den Fall, daß nur ein einziges Teilchen vorhanden ist, so daß etwa

$$N_s' = \delta_{s\sigma} \qquad (22)$$

In diesem Falle muß sich der die Wechselwirkungsenergie darstellende Operator auf Null reduzieren. In der Tat bleiben in der Summe $\sum_{\mu\nu rs}$ $N_r'-\delta_{r\sigma}$, der nach (22) stets verschwindet.

§ 2. Vergleich mit der Wellengleichung im Koordinatenraum.

Zur Rechtfertigung der Gleichung (21) für den Fall mehrerer Teilchen haben wir sie zu vergleichen mit der von Dirac auf Grund der Schrödingerschen Methode gegebenen Darstellung des Mehrkörperproblems mit symmetrischen Eigenfunktionen. Wir numerieren die Atome mit $k=1,2,\dots N$; die Schrödingersche Gleichung lautet dann

$$\sum_k\left\{-\frac{h^2}{8\pi^2\mu}\Delta_k\psi-\varepsilon V_0(r_k)\psi\right\}+\varepsilon^2\sum_{k<l}G(r_k,r_l)+\frac{h}{2\pi i}\frac{\partial\psi}{\partial t}=0, \qquad (23)$$

wo Δ_k den zu den Koordinaten des k-ten Teilchens gehörigen Laplaceschen Operator bezeichnet, während r_k den Koordinatenvektor desselben Punktes bedeutet, und $G(r_k,r_l)$ die zu den

Lagen der beiden durch k und l bezeichneten Teilchen gehörige Greensche Funktion bezeichnet, ~~während r_k den Koordinatenvektor desselben Punktes bedeutet, und $G(r_k,r_l)$~~

Die Gleichung (23) lösen wir durch einen Ansatz

$$\psi=\sum_{n_1,n_2\dots} b(n_1,n_2,\dots)\,u_{n_1}(r_1)\,u_{n_2}(r_2)\dots, \qquad (24)$$

wo die u_n wie im vorigen Paragraphen die Eigenfunktionen des ungestörten Einelektron~~en~~ seiner Argument sein. Einsetzen von (24) in (23) ergibt mit Rücksicht auf (6)

$$-\frac{h}{2\pi i}\frac{\partial b(n_1,n_2\dots)}{\partial t}=\sum_k E_{n_k} b(n_1,n_2,\dots)$$
$$+\varepsilon^2\sum_{k<l}\sum_{m_k m_l} A(m_k m_l|n_k n_l)\,b(n_1,n_2,\dots,m_k,\dots m_l,\dots) \qquad (25)$$

Da b in $n_1,n_2,\dots$ symmetrisch ist, können wir diese Größe auch als Funktion der Zahlen $N_1',N_2'\dots$ betrachten, wo N_r' die Anzahl derjenigen n_k bezeichnet, die gleich r sind; d.h. N_r' hat dieselbe Bedeutung wie im vorigen Paragraphen. Hierbei müssen wir, wie bei Dirac erläutert, wegen der Bedeutung von $|b|^2$ als Zustandswahrscheinlichkeit die Normierung ändern. Wir setzen mit Dirac

$$b(N_1',N_2',\dots)=\left(\frac{N'!}{N_1'!N_2'!\dots}\right)^{1/2} b(n_1,n_2,\dots) \qquad (26)$$

wo $N'=\sum_k N_r'$ ist und die $n_1,n_2,\dots$ eins der zu den bestimmten Zahlen $N_1',N_2'\dots$ gehörigen Wertsysteme der Zahlen n_k bezeichnen. Es wird nunmehr, wie leicht ersichtlich, auch bei zusammenfallenden Indizes

$$b(n_1,n_2,\dots n_{k-1},\overset{n}{\cancel{m}}_k,\overset{n}{\cancel{m}}_{k+1},\dots,n_{l-1},m_l,n_{l+1},\dots)$$
$$=e^{\frac{2\pi i}{h}(\Theta_{n_k}+\Theta_{n_l}-\Theta_{m_k}-\Theta_{m_l})}\left(\frac{N_1'!N_2'!\dots}{N_k'!}\right)^{1/2} b(N_1',N_2'\dots), \qquad (27)$$

wo $e^{\frac{2\pi i}{h}\Theta_\mu}$, wie vorhin, einen Operator bezeichnet, der, auf eine Funktion der zu Θ_μ gehörigen Größe N_μ' angewandt, die Zahl N_μ' durch $N_\mu'-1$ ersetzt. Damit erhalten wir

$$-\frac{h}{2\pi i}\frac{\partial b(N_1',N_2'\cdots)}{\partial t} = b(N_1',N_2',\cdots)\cdot\sum_r N_r' E_r$$
$$+\varepsilon^2\sum_{k<l}\sum_{m_k m_l} A(m_k m_l | n_k n_l)\cdot(N_1'!N_2'!\cdots)^{-1/2}$$
$$\cdot e^{\frac{2\pi i}{h}(\Theta_{n_k}+\Theta_{n_l}-\Theta_{m_k}-\Theta_{m_l})}(N_1'!N_2'!\cdots)^{1/2}\, b(N_1',N_2'\cdots) \qquad (28)$$

Wir führen die Summation so aus, daß wir zunächst bei festen k, l die Zahlen $\mu = m_k$, $\nu = m_l$ alle Werte $1, 2, \cdots$ durchlaufen lassen; so entsteht also eine Summe $\sum_{\mu\nu} A(\mu\nu | n_k n_l)$. Die Summation nach k und l geschieht so, daß wir diejenigen Glieder zusammenfassen, in ~~einem~~ denen $n_k = r$ und $n = s$ wird. Bei einem bestimmten Wertsystem $n_1, n_2, \cdots$ sind dies $N_r' N_s'$ Glieder, wenn $r \neq s$, und $\frac{1}{2} N_r'(N_r'-1)$, wenn $r = s$; aber im ersten Falle muß noch ein Faktor $\frac{1}{2}$ hinzugefügt werden, weil sonst bei der Summation nach r und s jedes Elektronenpaar doppelt gezählt werden würde. Also bekommen wir endlich

$$-\frac{h}{2\pi i}\frac{\partial b(N_1',N_2',\cdots)}{\partial t} = b(N_1',N_2'\cdots)\sum_r N_r E_r$$
$$+\frac{\varepsilon^2}{2}\sum_{\mu\nu r s} A(\mu\nu | rs)\, N_r'(N_s'-\delta_{rs})\{N_1'!N_2'!\cdots\}^{-1/2} e^{\frac{2\pi i}{h}(\Theta_r+\Theta_s-\Theta_\mu-\Theta_\nu)}$$
$$\{N_1'!N_2'!\cdots\}^{1/2}\, b(N_1' N_2'\cdots). \qquad (29)$$

Es ist aber

$$\{\prod_\varrho N_\varrho'!\}^{-1/2} e^{\frac{2\pi i}{h}(\Theta_r+\Theta_s-\Theta_\mu-\Theta_\nu)}\{\prod_\sigma N_\sigma'!\}^{1/2}$$
$$=\{\prod_\varrho N_\varrho'!\}^{-1/2}\{\prod_\sigma(N_\sigma'-\delta_{\sigma r}-\delta_{\sigma s}+\delta_{\sigma\mu}+\delta_{\sigma\nu})!\}^{1/2} e^{\frac{2\pi i}{h}(\Theta_r+\Theta_s-\Theta_\mu-\Theta_\nu)}$$
$$=\{N_r'(N_s'-\delta_{rs})\}^{1/2}\{N_\nu'+1-\delta_{\nu s}-\delta_{\nu r}\}^{1/2}(N_\mu'+1+\delta_{\mu\nu}$$
$$-\delta_{\mu s}-\delta_{\mu r}\}^{1/2} e^{\frac{2\pi i}{h}(\Theta_r+\Theta_s-\Theta_\mu-\Theta_\nu)} \qquad (30)$$

folglich ist (29) dieselbe Gleichung wie (21), oder

$$\varphi(N_1',N_2',\cdots) = b(N_1',N_2'\cdots) \qquad (31)$$

§ 3. Zur Frage der Reihenfolge der Faktoren in dem Energieausdruck. Wir wollen hier einige Bemerkungen über die

Frage der Reihenfolge der Faktoren in dem Ausdruck (12) für die Gesamtenergie des Wellensystems machen und betrachten zu diesem Zwecke wieder das Variationsprinzip (4). Es sei ~~η(~~ $\eta(x'-x'') = \eta(x''-x')$ eine (singuläre) Matrix, die so definiert ist, daß

$$\int \eta(x-x')\varphi(x')dv' = \Delta\varphi(x) \qquad (32)$$

Wir können dann schreiben

$$\int(\operatorname{grad}\varphi^*\operatorname{grad}\varphi)dv = -\iint\eta(x'-x'')\varphi^*(x')\varphi(x'')dv'dv'', \qquad (33)$$

und folglich

$$L = \int L\,dv = \frac{h}{4\pi i}\int\left\{\varphi^*(x)\frac{\partial\varphi(x)}{\partial t} - \frac{\partial\varphi^*(x)}{\partial t}\varphi(x)\right\}dv$$
$$-\frac{h^2}{8\pi^2\mu}\iint\eta(x'-x'')\varphi^*(x')\varphi(x'')dv'dv'' + \frac{\varepsilon^2}{2}\iint G(x',x'')\varphi^*(x')\varphi(x')$$
$$\times\varphi^*(x'')\varphi(x'')dv'dv''. \qquad (34)$$

Ebenso bekommen wir

$$E = -\frac{h^2}{8\pi^2\mu}\iint\eta(x'-x'')\varphi^*(x')\varphi(x'')dv'dv''$$
$$-\varepsilon\int V_0\varphi^*\varphi\,dv + \frac{\varepsilon^2}{2}\iint G(x',x'')\varphi^*(x')\varphi(x')\varphi^*(x'')\varphi(x'')dv'dv'' \qquad (35)$$

In dieser Form stellt (4) das Hamiltonsche Prinzip eines Systems von unendlich vielen Unbekannten dar, bei dem $\frac{h}{2\pi i}\varphi(x)$ und $\varphi^*(x)$ für alle Punkte x des Raumes die Rolle der kanonischen Veränderlichen spielen. Durch Grenzübergang von dem Fall endlich vieler Unbekannten werden wir zu den folgenden Quantenbedingungen geführt:

$$\varphi(x')\varphi^\dagger(x'') - \varphi^\dagger(x'')\varphi(x') = \delta(x'-x''), \qquad (36)$$

wo $\delta(x'-x'')$ der von Dirac ~~ein~~ eingeführten symbolischen Funktion entspricht, und, wenn $F(x)$ eine beliebige Funktion von x bedeutet, definiert wird durch

$$F(x) = \int F(x')\delta(x'-x)dv'. \qquad (37)$$

In der Tat können wir, auf Grund von (36), wie leicht zu zeigen ist, die Gleichungen für φ und φ^*, deren klassisches Analogon aus dem Variationsprinzip~~s~~ (4) folgt, in der folgenden Form darstellen:

$$\left.\begin{aligned}\frac{h}{2\pi i}\frac{\partial\varphi(\mathfrak{r})}{\partial t} &= E\varphi(\mathfrak{r})-\varphi(\mathfrak{r})E,\\ \frac{h}{2\pi i}\frac{\partial\varphi^{\dagger}(\mathfrak{r})}{\partial t} &= E\varphi^{\dagger}(\mathfrak{r})-\varphi^{\dagger}(\mathfrak{r})E,\end{aligned}\right\}\qquad(38)$$

so daß die Forderungen der Quantentheorie wirklich erfüllt sind. ~~Wei~~ Wenn wir (7) durch die entsprechenden q-Zahlengleichungen

$$\left.\begin{aligned}\varphi(\mathfrak{r}) &= \sum_s b_s(t)\,u_s(\mathfrak{r})\\ \varphi^{\dagger}(\mathfrak{r}) &= \sum_s b_s^{\dagger}(t)\,u_s(\mathfrak{r})\end{aligned}\right\}\qquad(7')$$

ersetzen, so können wir direkt von (14) zu (36) übergehen, und umgekehrt[1].

In (35) haben wir diejenige Reihenfolge der Faktoren gewählt, die sich naturgemäß ergibt, wenn wir das Gleichungssystem (1) in die Quantentheorie übernehmen. Dies ist aber nicht richtig, wie ein Vergleich mit (12) ~~zu~~ auf Grund von (7) zeigt, sondern in dem letzten Gliede von (35) müssen beim Ersatz der c-Zahlgrößen φ, φ^* durch q-Zahlen $\varphi, \varphi^{\dagger}$ die beiden mittleren Faktoren $\varphi(\mathfrak{r}')$ und $\varphi^{\dagger}(\mathfrak{r}'')$ vertauscht werden. Nun ist nach (36)

$$\varphi(\mathfrak{r}')\,\varphi^{\dagger}(\mathfrak{r}'') = \varphi^{\dagger}(\mathfrak{r}'')\,\varphi(\mathfrak{r}') + \delta(\mathfrak{r}'-\mathfrak{r}'').$$

Also folgt, wenn wir das in (35) auftretende E mit E' bezeichnen und E für den richtigen Energieausdruck schreiben:

$$E = E' - \frac{\varepsilon^2}{2}\iint G(\mathfrak{r}',\mathfrak{r}'')\,\varphi^{\dagger}(\mathfrak{r}')\,\varphi(\mathfrak{r}'')\,\delta(\mathfrak{r}'-\mathfrak{r}'')\,dv'\,dv''$$

oder, wenn wir einen Augenblick von der Singularität von $G(\mathfrak{r}',\mathfrak{r}'')$ für $\mathfrak{r}'=\mathfrak{r}''$ ~~an~~ absehen:

$$E = E' - \frac{\varepsilon^2}{2}\int G(\mathfrak{r},\mathfrak{r})\,\varphi^{\dagger}(\mathfrak{r})\,\varphi(\mathfrak{r})\,dv$$

Diese Ausdruck hat eine anschauliche Bedeutung.

Klassisch ist nämlich die Wechselwirkungsenergie eines Systems von Massenpunkten, wenn zwischen je zweien eine Wechselwirkungsenergie $\varepsilon^2 G(\mathfrak{r}_k,\mathfrak{r}_l)$ besteht, gegeben durch

$$\frac{\varepsilon^2}{2}\sum_{k\neq l} G(\mathfrak{r}_k,\mathfrak{r}_l) = \frac{\varepsilon^2}{2}\sum_{k,l} G(\mathfrak{r}_k,\mathfrak{r}_l) - \frac{\varepsilon^2}{2}\sum_{k} G(\mathfrak{r}_k,\mathfrak{r}_l).$$

Wenn nun im Volumenelement dv' gerade $N(\mathfrak{r}')\,dv' = \varphi^{\dagger}(\mathfrak{r}')\,\varphi(\mathfrak{r}')\,dv'$ dieser Massenpunkte liegen, so kann man die obige Summe ersetzen durch die Integrale:

$$\frac{\varepsilon^2}{2}\iint G(\mathfrak{r}',\mathfrak{r}'')\,N(\mathfrak{r}')\,N(\mathfrak{r}'')\,dv'\,dv'' - \frac{\varepsilon^2}{2}\int G(\mathfrak{r},\mathfrak{r})\,N(\mathfrak{r})\,dv \qquad (40)$$

Diejenige Ausdruck der Wechselwirkungsenergie, den wir oben benutzt haben, ist also in der Tat genau derjenige, den man bei sinngemäßer Analogisierung der klassischen Theorie erwarten muß. Die nicht kommutative Multiplikation der Q. Mech. macht es bemerkenswerterweise möglich, die Differenz eines zweifachen und eines einfachen Volumenintegrals in (40) durch ~~einz~~ einziges zweifaches Volumenintegral auszudrücken; dadurch wird es in der Quantenmechanik möglich, in analytisch einfacher Weise zum Ausdruck zu bringen, daß das „Eigenfeld" eines Elektrons auf dieses nicht in derselben Weise wirkt wie das „Außenfeld" — ein Unterschied, der innerhalb der klassischen Theorie sehr unbefriedigend und nur schwer exakt zu formulieren schien.

§4. Behandlung allgemeiner Wechselwirkungen zwischen den Teilchen.

Wir wollen nun dazu übergehen, die magnetischen Wechselwirkungen zwischen den Elektronen, soweit sie von der Größenordnung $1/c^2$ sind, mit zu berücksichtigen. Es müssen dann entsprechend auch die mit $1/c^2$ proportionalen Glieder in der Energie jedes einzelnen Teilchens berücksichtigt werden, die von der relativistischen Massenveränderlichkeit herrühren. Außerdem liefert die Retardierung der elektrischen (nicht der magnetischen) Kräfte bereits Energiekorrektionen der Größenordnung $1/c^2$; auf diese Korrektion kommen wir am Schluß zurück. Bevor wir das soeben gestellte Problem genauer betrachten, wollen wir uns klarmachen, daß die in den vorangehenden

---

[1] Ausführlicher wird dieser Zusammenhang der Gl. (7'), (14), (36) in einer nachfolgenden Arbeit von Jordan erörtert. Über die Verallgemeinerung dieser Gleichungen in der relativistisch invarianten Quantenelektrodynamik siehe eine demnächst erscheinende Arbeit von Pauli und Jordan.

$(a_\mu^{+}\varphi_\mu^{+} + b_\mu^{+}\varphi_\mu^{+})(\sum a_\mu \varphi_\mu + b_\mu \varphi_\mu)$

$=$

Paragraphen erzielten Ergebnisse eine sehr weitgehende Verallgemeinerung gestatten. Wir wollen nämlich ein Mehrkörperproblem betrachten, dessen Energie in korpuskulartheoretischer Darstellung die Gestalt

$$H = \sum_{k=1}^{N} H_1(p_k, q_k) + \frac{1}{2}\sum_{k,l=1}^{N}{}' H_2(p_k, q_k; p_l, q_l) \quad (41)$$

besitzt.

Hier kann die Energie $H_1(p_k, q_k)$ eines einzelnen Teilchens — durch p_k, q_k sollen sämtliche zum k-ten Teilchen gehörigen Impulse und Koordinaten angedeutet sein — ganz beliebig gewählt ~~ter~~ werden (jedoch für alle Teilchen in gleicher Weise); zwischen je zwei Massenpunkten soll dann eine gleichfalls ~~bel~~ beliebig zu wählende Wechselwirkung $H_2(p_k, q_k; p_l, q_l)$ bestehen, wobei

$$H_2(p_k, q_k; p_l, q_l) = H_2(p_l, q_l; p_k, q_k)$$

gewählt werden kann. Das Summenzeichen $\sum_{kl}{}'$ soll bedeuten, daß wieder Summanden mit $k = l$ ausschließen sind. Diese Voraussetzungen sind hinreichend, um die Anwendung der im vorangehenden entwickelten Überlegungen zu gestatten: Wir können unser Problem (41), wenn wir nur die symmetrischen Eigenfunktionen betrachten, ersetzen durch ein äquivalentes Wellenproblem mit gequantelten Wellen. Die Energie des Gesamtsystems kann dargestellt werden durch die Größe b_r^+, b_s bzw die mit ihnen durch (7') verbundenen gequantelten Wellenamplituden $\varphi^+(r)$, $\varphi(r)$; und zwar erhält man für diese Energie den Ausdruck

$$\left.\begin{aligned} E\ \cancel{H} = {} & \int dv'\,\varphi^+(r')\,H_1\!\left(\frac{h}{2\pi i}\frac{\partial}{\partial q'}, q'\right)\varphi(r') \quad E \\ & + \frac{1}{2}\iint dv'dv''\,\varphi^+(r')\,\varphi^+(r'')\,H_2\!\left(\frac{h}{2\pi i}\frac{\partial}{\partial q'}, q'; \frac{h}{2\pi i}\frac{\partial}{\partial q''}, q''\right)\varphi(r')\,\varphi(r'') \end{aligned}\right\}(42)$$

Zum Beweise dieser allgemeinen Formel sind wörtlich dieselben Betrachtungen auszuführen, die wir oben für den speziellen Fall benutzt haben.

Wenn wir nun die magnetische ~~Wirkung~~ Wechselwirkung der Elektronen zu bestimmen versuchen, so zeigt sich, daß die bisherigen Methoden keine eindeutige Vorschrift dafür liefern. Man kann klassisch diese

Wechselwirkungen, wie von Darwin¹⁾ gezeigt worden ist, aus einer Hamiltonfunktion

$$H_{kl} = -\frac{\varepsilon^2}{2c^2\mu^2}\frac{p_k p_l}{r_{kl}} \quad (43)$$

berechnen, wo p_k und p_l die Impulsvektoren der Elektronen sind und $r_{kl} = |r_k - r_l|$ ist. Bei den gewöhnlichen, auf die Matrizentheorie oder auf die Schrödingersche Wellengleichung im Koordinatenraum gegründeten Methoden ergibt sich nun für die exakte quantentheoretische Übersetzung dieser Gleichung eine Schwierigkeit daraus, daß in H_{kl} nichtvertauschbare Größen miteinander multipliziert sind, so daß die Forderung einer engen Analogie zur klassischen Formel (43) nicht ausreicht, um die Reihenfolge dieser Faktoren zu bestimmen.² Wir werden aber sehen, daß man bei der Darstellung der Elektronen durch gequantelte Wellen zwangläufig auf einen bestimmten Ansatz für die magnetische Wechselwirkungsenergie in (42) bringen, erhalten wir vermöge unseres allgemeinen Satzes die Möglichkeit, rückwärts auch die Form (41) ~~d~~ der Energie zu bestimmen und damit das exakte quantenmechanische Analogon der klassischen Formel (43) anzugeben.

Benutzen wir die $\varphi(r)$-Wellen zunächst in ihrer ungequantelten Form (also die Schrödingerschen Welle eines einzigen Elektrons), so werden wir, wenn $\mathfrak{I}(r)$ die elektrische Stromdichte ist, die magnetische Wechselwirkungsenergie gemäß dem Biot-Savartsche Gesetz (wie auch aus dem ~~voll~~ wellenmechanischen Variationsprinzip folgt) darstellen in der Form

$$-\frac{1}{2c^2}\iint \{\mathfrak{I}(r')\,\mathfrak{I}(r'')\}\,G(r, r'')\,dv'dv''. \quad (44)$$

In der für uns in Betracht kommenden Annäherung haben wir dabei

$$\mathfrak{I} = -\frac{\varepsilon}{2\mu}\frac{h}{2\pi i}\{\varphi^* \operatorname{grad}\varphi - \operatorname{grad}\varphi^*\,\varphi\} \quad (45)$$

zu setzen. Gehen wir nun von den c-Zahlengrößen φ^*, φ zu q-Zahlen φ^+, φ über, so haben wir den Ausdruck (44) in Rücksicht auf (45) zu ersetzen durch

$$H_1 = \frac{\varepsilon^2 h^2}{32\pi^2 c^2} \iint dv'\,dv''\, G(r',r'') \{ \varphi^\dagger(r')\varphi^\dagger(r'')\,\mathrm{grad}\,\varphi(r')\,\mathrm{grad}\,\varphi(r'')$$
$$- \varphi^\dagger(r')\,\mathrm{grad}\,\varphi^\dagger(r'')\,\mathrm{grad}\,\varphi(r')\varphi(r'')$$
$$- \mathrm{grad}\,\varphi^\dagger(r')\varphi^\dagger(r'')\,\varphi(r')\,\mathrm{grad}\,\varphi(r'')$$
$$+ \mathrm{grad}\,\varphi^\dagger(r')\,\mathrm{grad}\,\varphi^\dagger(r'')\,\varphi(r')\,\varphi(r'') \} \qquad (46)$$

Diesen Ausdruck können wir durch partielle Integrationen in die Gestalt (42) bringen:

$$H_1 = \frac{\varepsilon^2 h^2}{32\pi^2 c^2} \iint dv'\,dv''\,\varphi^\dagger(r')\varphi^\dagger(r'')\;\Omega\,\varphi(r')\varphi(r'') \qquad (47)$$

wobei der Operator Ω durch

$$\Omega = G\,\mathrm{grad}'\,\mathrm{grad}'' \ldots + \mathrm{grad}'G\,\mathrm{grad}'' \ldots + \mathrm{grad}''G\,\mathrm{grad}' \ldots$$
$$+ G\,\mathrm{grad}'\,\mathrm{grad}''\overset{G}{\cancel{-}} \ldots \cancel{+}$$
$$= 2\,(\,\mathrm{grad}'G\,\mathrm{grad}'' \ldots + \mathrm{grad}''G\,\mathrm{grad}' \ldots) \qquad (48)$$

gegeben ist. Dabei ist unter grad'; grad'' die Gradientenbildung an Funktionen von r' bzw r'' zu verstehen; daß die beiden Ausdrücke (48) für Ω übereinstimmen, findet man unter Berücksichtigung von[1]

$$\mathrm{grad}'\,\mathrm{grad}''\,G = -\mathrm{div}'\,\mathrm{grad}'\,G = 0\,. \qquad (49)$$

Aus (48) entnehmen wir dann das quantenmechanische Analogon der Formel (43) ~~in~~ in ~~die~~ der Gestalt

$$H_{kl} = -\frac{\varepsilon^2}{4c^2\mu^2}\left\{ p_k \frac{1}{r} p_l + p_l \frac{1}{r} p_k \right\} \qquad (50)$$

Wenngleich diese Formel abgeleitet wurde aus Betrachtungen, welche die Gültigkeit der Bose-Einsteinschen statt der Paulischen Statistik für Elektronen voraussetzen, so dürfte sie doch unabhängig von dieser Voraussetzung gültig sein; denn offenbar ist die Formel (50) von solcher Natur, daß ~~sie~~ keine Bezugnahme auf eine spezielle Wahl der Statistik mehr enthält.

Was endlich die mit $1/c$ proportionalen Effekte anbetrifft, die von der Retardierung der elektrischen Wechselwirkungen herrühren, so begnügen wir uns damit hervorzuheben, daß diese im Anschluß an die zitierte Darwinsche Untersuchungen in ganz derselben Weise behandelt werden können, wie die magnetischen Wechselwirkungen.

(Dem International Education Board haben wir zu danken für die Ermöglichung unserer Zusammenarbeit. Herrn Prof. Niels Bohr sind wir zu herzlichem Dank verpflichtet für viele lehrreiche Unterhaltungen über die Grundlagen der Quantentheorie.)

$$\varphi^\dagger(r')\,\varphi^\dagger(r'')\,\{G\,\mathrm{grad}'\,\mathrm{grad}'' +$$

$$G\,\mathrm{grad}'\,\mathrm{grad}''\,\varphi(r')\varphi(r'') = \overset{\mathrm{grad}}{\cancel{-}}\;\mathrm{grad}'G\,\cancel{\mathrm{grad}''}\,\varphi(r')\varphi(r'')$$
$$+ \mathrm{grad}'\,\mathrm{grad}''\,G\,\varphi(r')\varphi(r'')$$
$$= \cancel{G}\,\mathrm{grad} \ldots - \mathrm{grad}'\{G\,\mathrm{grad}'' \ldots\} - \mathrm{grad}'G\cdot\mathrm{grad}'' \ldots$$
$$+$$
$$+\{\mathrm{grad}''\{G\,\mathrm{grad}' \ldots\} + \mathrm{grad}'G\cdot\mathrm{grad}'' \ldots$$

$$H_{kl} = -\frac{\varepsilon^2}{4c^2\mu^2}\left\{ p_\mu \frac{1}{s^2} G\, p'_\nu + p'_\mu\, G\, p_\nu \right)$$
$$\{\alpha_\lambda p_\lambda\, G\, \alpha'_k p'_k + \alpha_k p'_k\, G\, \alpha_\lambda p_\lambda\}$$
$$\psi^*\psi^{*\prime}\{p_\lambda\, G\, \alpha_k p'_k\}\psi\,\psi'$$
$$(\alpha_\mu$$

[1] 1. In Gleichung (49) tritt die Größe $\mathrm{grad}'\,\mathrm{grad}''\,G$ auf, die natürlich von dem Operator $\mathrm{grad}'\,\mathrm{grad}''\,G\,\ldots$ in (48) wohl zu unterscheiden ist.

The wave equations of the electron.

C. G. Darwin

118 p. 654

1. In a recent paper Dirac* has brilliantly removed the defects before existing in the mechanics of the electron, and has shown how the phenomena usually called the "spinning electron" fit into place in the complete theory. He applies to the problem the method of q-numbers and, using non-commutative algebra, exhibits the properties of a free electron, and of an electron in a central field of electric force. In a second paper† he also discusses the rules of combination and the Zeemann effect. There are probably readers will will share the present writer's feeling that the methods of non-commutative algebra are harder to follow, and certainly much more difficult to invent, than are operations of types long familiar to analysis. Wherever it is possible to do so, it is surely better to present the theory in a mathematical form that dates from the time of Laplace and Legendre, if only because the details of the calculus have been so much more thoroughly explored. So the object of the present work is to take Dirac's system and treat it by the ordinary methods of wave calculus. The chief point of interest is perhaps the solution of the problem of the central field, which can be carried out exactly and leads to Sommerfeld's original formula for for the hydrogen levels. But it is also of ~~much~~ some interest to exhibit the relationship of the new theory to the previous equations which ~~are~~ were derived empirically ~~to the~~ by the present writer. It appears that those equations were an approximation to the new ones, derived by an approximate elimination of two of Dirac's four wave functions. We shall also review a few other points connected with the free electron, the emission of radiation from an atom and its magnetic moment, and shall outline a discussion of the Zeeman effect.

2. Dirac's guiding principle is that the "Hamiltonian equation" must be linear, and he adopts the form

$$p_0+\alpha_1 p_1+\alpha_2 p_2+\alpha_3 p_3+\alpha_4 mc=0,$$

where

$$\left.\begin{aligned} p_0 &= -\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{e}{c}V \\ p_1 &= \frac{h}{2\pi i}\frac{\partial}{\partial x}+\frac{e}{c}A_1, \text{etc} \end{aligned}\right\} \qquad (2,1)$$

V and v being scalar and vector potentials; while $\alpha_1 - \alpha_4$ are four four-rowed matrices obey the rules

$$p_1 \quad \frac{\partial}{\partial x}+i\frac{\partial}{\partial y} = e^{i\theta}\frac{\partial}{\partial r}+i\frac{e^{i\theta}}{r}\frac{\partial}{\partial\theta}$$

$$\frac{\partial}{\partial x}-i\frac{\partial}{\partial y} = e^{-i\theta}\frac{\partial}{\partial r}-i\frac{e^{-i\theta}}{r}\frac{\partial}{\partial\theta}.$$

$$\alpha_s^2 = 1 \qquad \alpha_s\alpha_t+\alpha_t\alpha_s=0$$

The α's are capable of an indefinite number of forms, and he gives rules for forming one set (though he does not write them out). The four-rowed matrices imply four wave functions which satisfy the simultaneous equations

$$\left.\begin{aligned}
&(p_0+mc)\psi_1+(p_1-ip_2)\psi_4+p_3\psi_3=0\\
&(p_0+mc)\psi_2+(p_1+ip_2)\psi_3-p_3\psi_4=0\\
&(p_0-mc)\psi_3+(p_1-ip_2)\psi_2+p_3\psi_1=0\\
&(p_0-mc)\psi_4+(p_1+ip_2)\psi_1-p_3\psi_2=0
\end{aligned}\right\} \begin{matrix}+\psi_3\\-\psi_4\\+\psi_1\\-\psi_2\end{matrix}\ (2,2)$$

We take these, then as our fundamental equations and discuss their solution employing only the ordinary methods of differential equations.

The equations are very unsymmetrical, and it is, of course, necessary first to show that they can be restored to their original form when axes are changed ~~for~~ or a relativity transformation is applied. The general formulae are complicated (being best expressed by four-dimensional Cayley parameters), but it is sufficient to verify the result for certain simpler transformations which can be imagined applied successively. This is also straightforward that we need merely give the results.

(1) Relativity Transformation

$$x=x' \quad y=y' \quad z=z'\cosh\beta+ct'\sinh\beta \qquad ct=ct'\cosh\beta+z'\sinh\beta$$

The equations are restored to their original form by

$$\psi_1'=\psi_1\cosh\frac{\beta}{2}+\psi_3\sinh\frac{\beta}{2}, \qquad \psi_2'=\psi_2\cosh\frac{\beta}{2}-\psi_4\sinh\frac{\beta}{2}$$

$$\psi_3'=\psi_3\cosh\frac{\beta}{2}+\psi_1\sinh\frac{\beta}{2}, \qquad \psi_4'=\psi_4\cosh\frac{\beta}{2}-\psi_2\sinh\frac{\beta}{2}.$$

(2) Rotation about z.

$$x=x'\cos\alpha-y'\sin\alpha, \quad y=y'\cos\alpha+x'\sin\alpha, \quad z=z', \quad t=t$$

Then $\psi_1'=\psi_1e^{i\alpha/2}$, $\psi_2'=\psi_2e^{-i\alpha/2}$, $\psi_3'=\psi_3e^{i\alpha/2}$, $\psi_4'=\psi_4e^{-i\alpha/2}$

(3) Rotation about y.

$\alpha=2\pi$ $\psi_1'=-\psi_1$ $\psi_2'=-\psi_2$ $\psi_3'=-\psi_3$ $\psi_4'=-\psi_4$

$$z=z'\cos\alpha-x'\sin\alpha, \quad x=x'\cos\alpha+z'\sin\alpha, \quad y=y' \quad t=t'$$

$$\psi_1'=\psi_1\cos\alpha/2+\psi_2\sin\frac{\alpha}{2}, \qquad \psi_2'=\psi_2\cos\frac{\alpha}{2}-\psi_1\sin\frac{\alpha}{2}$$

$$\psi_3'=\psi_3\cos\frac{\alpha}{2}+\psi_4\sin\frac{\alpha}{2}, \qquad \psi_4'=\psi_4\cos\frac{\alpha}{2}-\psi_3\sin\frac{\alpha}{2}.$$

These three transformations can build a group which represents any relativity

transformation, and so the invariance is proved.

It is of some interest to consider this invariance a little further. The whole theory of general relativity is based on the idea of invariance of form, and here we have a system invariant in fact but not in form. Should it not be possible to give it formal invariance as well, and would not that be the right way to express our equations? It is so possible, but it is not hard to show that it requires no less than 16 quantities to do it,* viz., two scalars, two four-vectors and one six-vector, and even so each will have a real and imaginary part, so that we may say that 32 quantities are required! It seems quite preposterous to think that a single electron should require 32 equations to express its behaviour, and, moreover, these will involve a large number of arbitrary inter-relations of no influence on the four quantities which are actually sufficient to describe it. Now the relativity theory is based on ~~the~~ nothing but the idea of invariance, and develops from it the conception of tensors as a matter of nessity; and it is rather disconcerting to find that apparently something has slipped through the net,† so that physical quantities exist which it would be, to say the least, very artificial and inconvenient to express as tensors. It does not seem possible to make anything further out of the matter until it has developed more, and we shall be content with one observation. Unlike the electromagnetic equations, our wave equations are homogeneous, so that there is no external quantity, like the electric current, etc., which could, so to speak, anchor them down in form to a definite set of directions. Now, there ought to be something of the kind because of the electromagnetic field of the electron, which in classical theory is made responsible for its mass. So we may perhaps conclude that it is not to be expected that our equations will attain a final form until the terms in mc are eliminated, that is, until we know how to do in the quantum theory a calculation like that which gives electromagnetic mass in the classical.

In my earlier paper a similar question arose and was much more easily resolved.

---

* 'Nature', vol 119, p. 282 (1927) 116 p.227.

* We can express the equation as a group of 16 in

$$\alpha\psi_1+\beta\psi_2+\gamma\psi_3+\delta\psi_4, \quad \alpha\psi_4+\beta\psi_3+\gamma\psi_2+\delta\psi_1, \ \text{etc.},$$

with α β γ δ arbitrary constants and can throw these into tensor form.

† Our equations (2,2) do not, of course, include gravitation, and this may be the hole in the net. But if gravitation were included, we should presumably be forced to introduce the tensor form, involving 16 complex or 32 real quantities, and this does not seem physically very plausible.

$2Y_0 = Z+W$
$2X_0 = Z-W$
$X+Y = 2U$
$Y-X = 2V$

$p_0W + mcZ = (p.X) = pU$
$p_0Z + mcW = pV$

In that work there were only two functions instead of the four here, and it was an easy matter to throw them into space-vector form, though it involved having four equations instead of two with a corresponding arbitraryiness in the solution. It appeared reasonable to make the step from two to four, and so to gain the advantage of vector notation, but to expand from four to sixteen is a different matter, and suggests that even in the simpler case the expansion is rather artificial. Nevertheless, it is not without interest to exhibit Dirac's equations in the form of space-vectors, without bringing in the time as part of the vectorial system. This can be done by a method similar to the substitution (5.1) of my paper. Take two vectors X, Y and two scalars X_0, Y_0 and write p_1, p_2, p_3 as a vector p. The equations

$$\left\{\begin{aligned}(p_0+mc)\,\cancel{X_0}Y_0 &= (p,X)\\ (p_0+mc)Y + p\cdot X_0 &= [p,X]\\ -(p_0-mc)X_0 &= (p,Y)\\ (p_0-mc)X - pY_0 &= -[p,Y]\end{aligned}\right\}$$

$X_i + iY_i = Z_i$

$\operatorname{div} Z + \frac{i}{c}\frac{\partial Z_0}{\partial t} = mcZ_0 = 0$ (2,3)

$\operatorname{div} Z \quad \frac{\partial Z_3}{\partial t} - \frac{\partial Z_2}{\partial x} + \frac{\partial Z_1}{\partial y} - \frac{\partial Z_0}{\partial z} = 0$

can be combined together in pairs according to either of the following alternative schemes so as to give Dirac's equations;—

$$\left.\begin{aligned}\psi_1 &= Y_3 - iY_0\\ \psi_2 &= Y_1 + iY_2\\ \psi_3 &= iX_3 \mp X_0\\ \psi_4 &= iX_1 - X_2\end{aligned}\right\} \qquad \left.\begin{aligned}\psi_1 &= -Y_1 + iY_2\\ \psi_2 &= Y_3 + iY_0\\ \psi_3 &= -iX_1 - X_2\\ \psi_4 &= iX_3 + X_0\end{aligned}\right\}$$

$Y_0 = -Y_0$
$\psi_1 + \psi_3 = iZ_3 + Z_0$
$\psi_2 - \psi_4 = Z_2 - iZ_1$

$$\left\{\begin{aligned}\operatorname{Div} Z &= mcZ_0\\ \operatorname{Rot} Z &= mcZ_0\end{aligned}\right.$$

$$\left\{\begin{aligned}Z^i_{.i} &= kZ^0_0\\ Z^i_{.k} &= kZ^i_i\end{aligned}\right.$$

So we might regard (2,3) as the primitive equations giving Dirac's twice over. Now the operation $p_1 = \frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_1$ may be likened to a "covariant differentiation" (say, by the introduction of a fifth dimension in the manner of Klein*), and in this sense $[p,X]$ may be called curl X and (p,X), div X. In the same sense p_0 is a time differentiation, and we see that (2,3) bear a rather striking resemblance to the electromagnetic equations with X and Y for E and H, and X_0, Y_0 playing a role ~~taken~~ akin to electric and magnetic density. It does not seem possible to push this rather loose analogy farther at present, and again we have a hint as to the reason, because there is no electromagnetic analogue to the terms in mc; and this will only be supplied when we know how to calculate electromagnetic mass in the quantum theory.

* Klein; Zs. f. Phys. 37 p. 895

$mcV - p\cdot W = -[p,V]$
$p_0V + mcV + pZ = [p\,U]$

3. The equations (2,2) are sufficient to determine the levels of any system; but that is not enough, for we also require to know the rules of combination, and for this we must have the extension of the "electric density" of Schrödinger's theory to the present case. In order to find the radiation of an atom we combine atom and radiation field into a single system, expressing the equations all together most conveniently by means of a variational principle.

The variational function is easy to construct. Multiply the equations (2,2) by $\delta\psi_1^*, \delta\psi_2^*, \delta\psi_3^*, \delta\psi_4^*$ ~~re~~ respectively and add them together. Add on their conjugate and integrate over all space and time. Integrating by parts, this can be written as the variation of a single function. To this we must add $\frac{1}{8\pi}(H^2-E^2)$, integrated over space and time, in order to give the dynamical effect of the radiation. Then if we express H and E in terms of the potentials, and vary V and the A's as well as the ψ's and ψ^*'s, we shall obtain equations for V and A's as well as (2,2) and their conjugates, and shall thus have the radiation completely linked with the material ~~system~~ motion. In carrying out this process one point is to be noticed. The equations (2,2) are homogeneous and so do not fix the magnitudes of the ψ's. The equation for V is not so, ~~that~~ but determines V as proportional to a quadratic expression in the ψ's. We naturally adjust this by normalising the ψ's so that the total charge shall be $-e$, the charge of an electron. It is more convenient to normalise the ψ's to unity, and therefore the terms $\frac{1}{8\pi}(H^2-E^2)$ must be multiplied by a suitable constant. To save writing out many formula twice, we shall anticipate the knowledge of this constant; it is $-1/c$. Proceeding in this way we arrive at the variational function.

$$\begin{aligned}S = \iiiint dx\,dy\,dz\,dt\Big\{\tfrac{1}{2}\tfrac{h}{2\pi i}\Big[&\psi_1^*\Big(-\tfrac{1}{c}\tfrac{\partial\psi_1}{\partial t} + \Big(\tfrac{\partial}{\partial x} - i\tfrac{\partial}{\partial y}\Big)\psi_4 + \tfrac{\partial}{\partial z}\psi_3\Big)\\ &+\psi_2^*\Big(-\tfrac{1}{c}\tfrac{\partial\psi_2}{\partial t} + \Big(\tfrac{\partial}{\partial x} + i\tfrac{\partial}{\partial y}\Big)\psi_3 - \tfrac{\partial}{\partial z}\psi_4\Big)\\ &+\psi_3^*\Big(-\tfrac{1}{c}\tfrac{\partial\psi_3}{\partial t} + \Big(\tfrac{\partial}{\partial x} - i\tfrac{\partial}{\partial y}\Big)\psi_2 - \tfrac{\partial}{\partial z}\psi_1\Big)\\ &+\psi_4^*\Big(-\tfrac{1}{c}\tfrac{\partial\psi_4}{\partial t} + \Big(\tfrac{\partial}{\partial x} + i\tfrac{\partial}{\partial y}\Big)\psi_1 - \tfrac{\partial}{\partial z}\psi_2\Big)\Big] + \text{the conjugate of the proceeding.}\\ &+mc\,(\psi_1\psi_1^* + \psi_2\psi_2^* - \psi_3\psi_3^* - \psi_4\psi_4^*) + \tfrac{eV}{c}\sum_1^4\psi_\lambda\psi_\lambda^*\\ &+\tfrac{eA_2}{c}\,(-i\psi_1^*\psi_4 + i\psi_2^*\psi_3 - i\psi_3^*\psi_2 + i\psi_4^*\psi_1)\end{aligned}$$

$\left(\frac{h}{4\pi i}\left(\frac{1}{c}\frac{\partial\rho}{\partial t} + \operatorname{div} \mathrm{I}\right) + mc\,F + \frac{eV}{c^2}\rho + \frac{eA\,\mathrm{I}}{c^2}\right)$

$\left(-\frac{1}{2c^2}G\right)$

$\left(\frac{c^2h}{4\pi i}\left(\frac{1}{c}\frac{\partial\rho}{\partial t} + \operatorname{div}\mathrm{I}\right) + mc^2F + eV\rho + eA\,\mathrm{I} - \frac{1}{2}G\right) = 0$

$$\psi_1\psi_4^*+\psi_4^*\psi_1=(X_1+iY_1)(X_4-iY_4)+(X_1-iY_1)(X_4+iY_4)=X_1X_4+Y_1Y_4$$

$$+\frac{eA_3}{c}(\psi_1^*\psi_3-\psi_2^*\psi_4+\psi_3^*\psi_1-\psi_4^*\psi_2)-\frac{1}{8\pi c}\left[(\mathrm{curl}\,A)^2-\left(\frac{1}{c}\frac{\partial A}{\partial t}+\mathrm{grad}\,V\right)^2\right]\Big\} \qquad (3,1)$$

If we vary this function it is easy to see that we obtain (2,2) and their conjugates and also

$$\left.\begin{aligned}\frac{1}{4\pi}\Box V&=e\sum_1^4\psi_\lambda^*\psi_\lambda\\ \frac{1}{4\pi}\Box A_1&=-e\{\psi_1^*\psi_4+\psi_2^*\psi_3+\psi_3^*\psi_2+\psi_4^*\psi_1\}\end{aligned}\right\} \qquad (3,2)$$

etc where $\Box$ signifies $\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{c}\frac{\partial^2}{\partial t^2}$. We conclude that the electromagnetic effect of the electron can be represented by taking density ρ and the current densities j_1, j_2, j_3, where

$\psi_1=H$

$$\left.\begin{aligned}\rho&=-e(\psi_1^*\psi_1+\psi_2^*\psi_2+\psi_3^*\psi_3+\psi_4^*\psi_4)\\ j_1&=ce(\psi_1^*\psi_4+\psi_2^*\psi_3+\psi_3^*\psi_2+\psi_4^*\psi_1)\\ j_2&=ce(-i\psi_1^*\psi_4+i\psi_2^*\psi_3-i\psi_3^*\psi_2+i\psi_4^*\psi_1)\\ j_3&=ce(\psi_1^*\psi_3-\psi_2^*\psi_4+\psi_3^*\psi_1-\psi_4^*\psi_2)\end{aligned}\right\} \qquad (3,3)$$

provided that the ψ's are normalised by the rule

$$\iiint\sum_{\lambda=1}^4\psi_\lambda\psi_\lambda^*\,dx\,dy\,dz=1 \qquad (3,4)$$

Since S is invariant for relativity transformations, ρ and j will be covariant for such transformations, and this can also be easily verified by applying the transformation of §2. in turn.

It should be observed that ρ and j satisfy the equation of continuity

$$\frac{\partial\rho}{\partial t}+\mathrm{div}\,j=0 \qquad (3,5)$$

as may be directly verified with the use of (2,2). That (3,5) should be verified is in a sense the starting point of Dirac's argument. For if it had not been so spontaneously, we should have been compelled to force it by introducing into S a term $\left(\frac{1}{c}\frac{\partial V}{\partial t}+\mathrm{grad}\,A\right)F$ with F undetermined. The result would give extra terms in ρ and j involving $\frac{\partial F}{\partial t}$ and grad F, and the condition $\frac{1}{c}\frac{\partial V}{\partial t}+\mathrm{grad}\,A=0$ would then fix F. In general it would involve V and A, and therefore ρ and j would do so too, It was the presence of such terms in Klein's† evaluation of density and current that was objectionable and that led Dirac to his new equations.

To complete the rules for calculating intensities we have to break up ρ and j into terms corresponding to pairs of states. This can be done in the manner of Klein, but perhaps the following picture, though very incomplete† may make the process clearer, and may show under what conditions we expect to get line spectra with definite intensities; it is applicable to any system. Imagine that we have an assembly of atoms in an enclosure. The equations (2,2) and (3,2) together with appropriate boundary conditions, will describe the state of affairs. Thermal equilibrium will be produced, with the accompanying black radiation, and the equations will be quite insoluble, because in solving (2,2) the electromagnetic fields, themselves determined by (2,2) will not be small. At any instant of time we can imagine the state expressed by an expansion in proper functions, and the average values of the coefficients will be determined by the appropriate statistics — in particular, states of nearly equal energy will have equal average coefficients. Now if the enclosing barrier is suddenly removed, the radiation before present will spread away with the speed of light and the matter will be left only under the influence of any existing permanent electromag. forces. The problem is now soluble by approximations, first solving (2,2) for the ψ's, neglecting the radiation, and then substituting the values found in (3,2) to give the radiation. If ψ_λ is initially $\sum_p a_p\psi_\lambda^p$, the first process gives $\psi_\lambda=\sum_p a_p\psi_\lambda^p e^{-i\frac{2\pi}{h}W_pt}$. Next we form

$$\rho=\sum_p\sum_q\sum_{\lambda=1}^4 a_pa_q^*\psi_\lambda^p\psi_\lambda^{q*}e^{i\frac{2\pi}{h}(W_q-W_p)t},$$

and similarly values for the j's. Substituting these in (3,2) we can evaluate the electric force at a distant point, and it will evidently consist of a sum of periodic terms corresponding to the spectrum lines given by W_p-W_q. The process is exactly that introduced by Klein, only a little more definite in that no appeal is made to the correspondence principle.

4. We shall next exhibit the relationship of Dirac's equations to previous theories, and shall show that the latter are successive approximations to (2,2). The guiding principle in this is the fact that of the four ψ's, ψ_3 and ψ_4 are very much larger than ψ_1, ψ_2, since this leads to a method of approximation. We shall treat of the case of stationary states of an electron in an atom.

The first approximation leads to Schrödinger's equation in both ψ_3 and ψ_4

† 41 407

† The incompleteness lies chiefly in the fact that no distinction is made between one atom and many atoms.

independently; both in doing this it must not be forgotten that that equation is not even approximately right as far as concerns the effects of magnetic fields. We must therefore restrict ourselves to equations in which $p_1 = \frac{h}{2\pi i}\frac{\partial}{\partial x}$ without a vector potential, so that $p_1 p_2 = p_2 p_1$, etc. Starting with complete neglect of ψ_1, ψ_2 we see that ψ_3, ψ_4, and therefore ψ_1, ψ_2 also, are proportional to $e^{-i\frac{2\pi}{h}mc^2t}$. Hence $(p_0 + mc)\psi_1$ is nearly equal to $2mc\psi_1$, and so we take

$$\left.\begin{aligned}\psi_1 &= -\frac{h}{2\pi i}\frac{1}{2mc}\left[\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\psi_4 + \frac{\partial}{\partial z}\psi_3\right]\\ \psi_2 &= -\frac{h}{2\pi i}\frac{1}{2mc}\left[\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\psi_3 - \frac{\partial}{\partial z}\psi_4\right]\end{aligned}\right\} \qquad (4,1)$$

Substituting in the third equation we get

$$\left(-mc + \frac{eV}{c} - \frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}\right)\psi_3 = \left(\frac{h}{2\pi i}\right)^2 \frac{1}{2mc}\Delta\psi_3$$

The same equation holds for ψ_4, so that we simply have Schrödinger's equation twice over.

In the second approximation, following Dirac's §4, we form exact second order equations in ψ_3, ψ_4, and from these we eliminate $\psi_1\ \psi_2$ by means of (4,1). In (2,2) operate the third equation by $(mc + p_0)$, the second by $-(p_1 - ip_2)$, the first by $-p_3$ and add. We have

$$p_0p_1 - p_1p_0 = -\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}\frac{eA_1}{c} - \frac{h}{2\pi i}\frac{\partial}{\partial x}\frac{eV}{c} = \frac{e}{c}\frac{h}{2\pi i}E_1$$

$$p_1p_2 - p_2p_1 = \frac{h}{2\pi i}\frac{\partial}{\partial x}\frac{eA_2}{c} - \frac{h}{2\pi i}\frac{\partial}{\partial y}\frac{eA_1}{c} = \frac{e}{c}\frac{h}{2\pi i}H_3$$

and using these and similar relations we get

$$\left[-m^2c^2 + p_0^2 - p_1^2 - p_2^2 - p_3^2 - \frac{e}{c}\frac{h}{2\pi}H_3\right]\psi_3 - \frac{e}{c}\frac{h}{2\pi}(H_1 - iH_2)\psi_4$$
$$-\frac{e}{c}\frac{h}{2\pi}i\left[(E_1 - iE_2)\psi_2 + E_3\psi_1\right] = 0$$

Now we substitute for ψ_1, ψ_2 from (4.1). These equations ought now to contain terms involving the A's, but as they are to be multiplied by E's, these may be neglected. We have

$$\left(D - \frac{2\pi e}{ch}H_3\right)\psi_3 - \frac{2\pi e}{ch}(H_1 - iH_2)\psi_4 + \frac{e}{2mc^2}\Big\{\left(E_1\frac{\partial}{\partial x} + E_2\frac{\partial}{\partial y} + E_3\frac{\partial}{\partial z}\right)\psi_3$$
$$+ i\left(E_2\frac{\partial}{\partial z} - E_3\frac{\partial}{\partial y}\right)\psi_4 + \left(E_3\frac{\partial}{\partial x} - E_1\frac{\partial}{\partial z}\right)\psi_4 + i\left(E_1\frac{\partial}{\partial y} - E_2\frac{\partial}{\partial x}\right)\psi_3\Big\} = 0 \qquad (4,2)$$

where $D = \left(\frac{2\pi}{h}\right)^2(-m^2c^2 + p_0^2 - p_1^2 - p_2^2 - p_3^2)$ is the same quantity as in my paper.

Similarly

$$\left(D + \frac{2\pi e}{ch}H_3\right)\psi_4 - \frac{2\pi e}{ch}(H_1 + iH_2)\psi_3 + \frac{e}{2mc^2}\Big\{\left(E_1\frac{\partial}{\partial x} + E_2\frac{\partial}{\partial y} + E_3\frac{\partial}{\partial z}\right)\psi_4 + i\left(E_2\frac{\partial}{\partial z} - E_3\frac{\partial}{\partial y}\right)\psi_3$$
$$-\left(E_3\frac{\partial}{\partial x} - E_1\frac{\partial}{\partial z}\right)\psi_3 - i\left(E_1\frac{\partial}{\partial y} - E_2\frac{\partial}{\partial x}\right)\psi_4\Big\} = 0 \qquad (4.3)$$

Apart from the terms in

$$P \equiv E_1\frac{\partial}{\partial x} + E_2\frac{\partial}{\partial y} + E_3\frac{\partial}{\partial z}, \qquad (4,4)$$

these are identical with the equations of my paper, provided that we identify ψ_3 with f, ψ_4 with $-g$. The extra terms in P ~~specify~~ rectify one of the earlier defects, for with ~~my~~ equations the s-levels of hydrogen fell in the wrong place though all others were correct. Now when the approximative method is used, the new terms affect the levels by an amount depending on $\int_0^\infty f(r)\,[P\cdot f(r)]\,r^2dr$ and for hydrogen $E_1 = ex/r^3$ so that $P = \frac{e}{r^2}\frac{d}{dr}$ and the integral depends on $[f(0)]^2$. Since $f(0)$ vanishes unless $k = 0$, all levels other than s-levels are unaffected, and a more detailed calculation shows that the s-levels now fall in the right place.

The formulae for intensities are also the same to a first approximation, because to this degree $ff^* + gg^*$ is the same as $\sum_2^4 \psi_\lambda\psi_\lambda^*$. It is of more interest to consider the formula for magnetic moment. For this we take all the ψ's in (3.3) as belonging to a single state, so that the time disappears and we have $\operatorname{div} j = 0$. With this condition and the condition $\iiint j\,dx\,dy\,dz = 0$ (which holds because there is no progressive current in a stationary ~~state~~ atom) the magnetic moment can be seen to have as first component

$$\mu_1 = \frac{1}{2c}\iiint(yj_3 - zj_2)\,dx\,dy\,dz.$$

Substitute (4,1) in (3,3) so as to obtain the approximations for j_2, j_3. After some partial integrations we find

$$\mu_1 = \frac{eh}{4\pi mc}\iiint\left[\frac{1}{2i}\left\{\psi_3R_1\psi_3^* + \psi_4R_1\psi_4^* - \psi_3^*R_1\psi_3 - \psi_4^*R_1\psi_4\right\} - (\psi_3\psi_4^* + \psi_3^*\psi_4)\right]dx\,dy\,dz \qquad (4.5)$$

where $R_1 = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}$. This is the value[†] given before, when we replace ψ_3, ψ_4 by f and $-g$. Similar methods apply for the other components. A cyclic change gives the first terms and the last are

$$-i(\psi_3\psi_4^* - \psi_3^*\psi_4) \quad \text{and} \quad +(\psi_4\psi_4^* - \psi_3\psi_3^*) \qquad (4,6)$$

respectively for μ_2 and μ_3.

[†] In §6 of my paper, by a blunder I took the mag. moment as $\frac{W}{\partial H}$ instead of $-\frac{\partial W}{\partial H}$. The expressions (6,17) there for the components of μ should have their signs changed. $x_1 = f$, $x_2 = if$, $x_3 = g$, $x_4 = ig$

Dirac's success in finding the accurate equations shows the great superiority of principle over the previous empirical method, but it is perhaps not without interest (at any rate to the present writer, who had projected but not begun such work) to consider whether the empirical method could have led by way of improved approximations to the accurate result. The most critical step in doing so had been made, though not quite rightly and for a wrong reason, in the replacement of $2mc$ by a time differential (it had been replaced by $2\left(-\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{eV}{c}\right)$ whereas only half should have been treated in this way). There also seemed nothing to prevent carrying out a higher approximation so as to make the hydrogen levels fall more exactly together. A further guide lay in the fact that the electric current must be a more primitive thing than the magnetic moment, and when the current is deduced from (4,5) (4,6), it has certain complicated small terms admitting of modification. On the other hand, the absence of the terms (4,4) would have caused trouble. On the whole, it seems not impossible that one might with much labour have arrived at some sort of elimination of Dirac's equations. Fortunately, he has made such work unnecessary.

5. The free motion of an electron calls for comment. In the equations (2,2) we now omit V and A altogether. Assume as solution

$$\psi_x = a_x \exp. i\frac{2\pi}{h}(px+qy+rz-Wt)$$

and on substituting we have a determinant which reduces to

$$(W^2-m^2c^2-p^2-q^2-r^2)^2=0.$$

We must therefore take $W=\sqrt{\{m^2c^2+p^2+q^2+r^2\}}$, which we shall call W_{pqr} for short. We can take a_3, a_4 as arbitrary, say A, B, and then have

$$a_1=-\frac{Ar+B(p-iq)}{mc+W_{pqr}} \qquad a_2=-\frac{A(p+iq)-Br}{mc+W_{pqr}}$$

From these we find that ρ and j are proportional to $|A^2|+|B^2|$, and that j is a vector along $p:q:r$.

An important point now arises if we consider the problem of from the point of view raised in a recent paper †. The motion of the electron is there regarded as a pure wave problem, the solution consisting in finding the way in which given arbitrary initial conditions are propagated. Suppose that we have arbitrary initial values of all four ψ's at every

point of space. We can submit them to Fourier analysis and have

$$\psi_x=\int a_x(p,q,r)\exp i\frac{2\pi}{h}(px+qy+rz)\,dp\,dq\,dr$$

The a_x's will have arbitrary values but this is impossible since we have just seen that a_1, a_2 are determinate in terms of a_3, a_4. There can hardly be a question that a complete theory will overcome this difficulty by admitting negative values of W, but we are evidently in contact with the question raised by Dirac in his § 1, connected with the possible changes of e into $-e$. At present this is unsolved, so we must be content to say that we are not entitled to assume completely arbitrary initial conditions, but may only take two of the four functions arbitrary.

To understand the physical meaning of the questions in free space, we want to be able to associate a given solution with the rectilinear motion of an electron with magnetic moment in a given direction. As long as we only deal with solutions of type (5,1) nothing can be said about the magnetic moment, because the wave fill all space uniformly and there is, therefore, no distant point left from which to observe it. In order to get a magnetic moment we must construct a wave packet. We may, for instance, assume that initially ψ_3 and ψ_4 contain a factor $\exp -\frac{1}{2\sigma^2}(x^2+y^2+z^2)$. These will fix ψ_1, ψ_2, and it is important to observe that in consequence of the differential inter-relation between them, the ratio of $\psi_1:\psi_2$ will vary in the different parts of the packet, so that the current j will ~~be~~ no longer be everywhere straight along the main direction of wave motion, and a magnetic moment becomes possible. The complete solution can be set down in Fourier integrals, but it does not seem possible to work them out. It is therefore pler to recur to the approximation of § 4, which shows that coefficients A and for ψ_3, ψ_4 will be associated with various moments. By consideration of the st term in (4,5), (4,6), we see that if A, B are respectively proportional to n) and $(-l-im)$ we shall have magnetic moment along the direction $l:m:n$. For high speeds of motion the approximation would fail, but so would the idea of agnetic moment. The same approximate method is adequate for the case of motion in a field and for the Stern-Gerlach effect. For the case of uniform field it is possible to find urate solutions of the equations, but they correspond to a quantised circular motion and are not of much interest, as they need to be combined into wave packets if any close relationship is to be seen with the velocity of the electron.

† 40 p. 258

$$F^{\mu\nu} \qquad (\mathrm{curl}\, A)^2 = (A^{\mu}_{\nu} - A^{\nu}_{\mu})$$

$$(A_{\mu\nu} - A_{\mu\nu})^2 - \quad -$$

118 Ionized. H-molecule, Wilson P. 635.

spinning Electron ノ存在ニヨッテ ↓ ノ問題ヲトクコト。

$$\delta\int \left(kJ^{\mu}_{.\mu} + J^{\mu}A_{\mu} + \cancel{A}\, F^{\mu\nu}F_{\mu\nu}\right)\sqrt{g}\, d\tau \quad {}^{+mcJ_0}$$

$$= 0$$

$$F^{\mu\nu}F_{\mu\nu} = \frac{1}{2}\{M^2 + \overline{M}^2\}$$

$$M = H + iE$$

$$J = \frac{1}{2}\{X^2 + \overline{X}^2\}$$

6. We now consider the energy levels for an electron in a central field, and in particular for hydrogen. Before proceeding to the solution, it will be well to discuss the question of the nomenclature of the various quantum states. Dirac points out that angular momentum is no longer an integral of the motion, but find a modified integral of a similar type. He ~~so~~ thus suppresses the use of k and uses a quantum number j. With the method of solution which we shall be using the dynamical meaning of quantum numbers goes very much into the background, and we are left only with integers defining the orders of spherical harmonics and other functions entering the solution. From this point of view the quantum number is only a convenient name associated with those functions — for example the quantum number m was adopted as $u + \frac{1}{2}$ in the earlier theory because a certain solution involved P_k^u and P_k^{u+1}. The fact that there exists a dynamical integral is then largely irrelevant — under special circumstances it might help in guessing a solution, but usually it merely reduces to an identity in the properties of a function found in some other way.

In view of these considerations, I have concluded, with some hesitation, that it is more convenient not to alter the notation in the way done by Dirac. The quantum numbers k, j, m retain their old classificatory, but not their dynamical significance; in this way any line of a spectrum can be described by the same numbers as were used before and doublets are classified like other multiplets. To define them more precisely we take m as given by Landé's g-formula, and j as the maximum positive value of m. For doublets it is therefore a half-number, and k may be either of the adjacent integers. We shall, in fact, see that every solution involves four different spherical harmonics, P_k^u, P_k^{u+1}, P_{k+1}^u, P_{k+1}^{u+1} and j and m are simply the averages of subscripts and of superscripts respectively. k, which is the subscript of ψs, has not the same symmetrical property as j and m, but is called a quantum number because it defines the rule of combination of levels and of the Zeeman effect. We adopt for it the numeration 0, 1, 2 … not 1, 2, 3, …, as this has proved most convenient in spectroscopy.

In ~~the~~ Dirac's new notation j can take on negative values, and it is at first sight tempting to follow this change. It would not be hard to invent negative spherical harmonics P^u_{-j}, which would make it possible to write all the solutions in a single form, but it lead to little simplification. For ~~this~~ j values do not run right through from positive to negative; they miss the value zero and it would be necessary to study combinations of like and of unlike signs separately, and this can be done just as well with only positive values and a second number k acting as a sort of plus or minus sign. Moreover, as we shall see, we can take full advantage of Dirac's method in discussing the radial functions. The following scheme shows the relationship between the values used here and by Dirac :—

		s	p		d	
Here	k	0	1	1	2	2
	j	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{2}$
Dirac	j	−1	1	−2	2	−3

7. In order to solve for the levels in a radial field of force, we put $p_0 = \frac{W}{c} + \frac{eV}{c}$ where V depends only on the radius, and we omit the vector potentials. The equations (2,2) now become :—

$$\left.\begin{aligned}
&\frac{2\pi i}{h}\left(\frac{W+eV}{c}+mc\right)\psi_1+\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\psi_4+\frac{\partial}{\partial z}\psi_3=0\\
&\frac{2\pi i}{h}\left(\frac{W+eV}{c}+mc\right)\psi_2+\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\psi_3-\frac{\partial}{\partial z}\psi_4=0\\
&\frac{2\pi i}{h}\left(\frac{W+eV}{c}-mc\right)\psi_3+\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\psi_2+\frac{\partial}{\partial z}\psi_1=0\\
&\frac{2\pi i}{h}\left(\frac{W+eV}{c}-mc\right)\psi_4+\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\psi_1-\frac{\partial}{\partial z}\psi_2=0
\end{aligned}\right\}\quad(7.1)$$

Following previous methods we try to express the four functions as spherical harmonics multiplied by radial functions.

We may first conveniently give certain formulae for spherical harmonics that are easily proved. We abreviate the notation by writing P^u_k for the whole harmonics, thus

$$P^u_k = (k-u)!\,\sin^u\theta\left(\frac{d}{d\cos\theta}\right)^{k+u}\frac{(\cos^2\theta-1)^k}{2^k k!}e^{iu\phi}$$

which exists for any positive integral values of k and for any integral values of u between ±k inclusive. Then, if f is any function of the radius we have :—

$$\frac{\partial}{\partial x}+i\frac{\partial}{\partial y} = \frac{\partial}{\partial(x+iy)} = \frac{1}{2}\frac{\partial}{\partial x}+\frac{1}{2i}\frac{\partial}{\partial y} = \frac{1}{2}\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)$$

$$\frac{\partial}{\partial x} = \quad r\sin\theta\, e^{i\phi} \quad z_1 = x+iy = x_1 \quad \frac{z_1+z_2}{2}=x \quad \frac{z_1-z_2}{2i}=y$$

$$z_2 = x-iy$$

$$\left\{\begin{aligned}
&\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)f\,P^u_k=\frac{1}{2k+1}\left\{\left(\frac{df}{dr}-\frac{k}{r}f\right)P^{u+1}_{k+1}-(k-u)(k-u-1)\left(\frac{df}{dr}+\frac{k+1}{r}f\right)P^{u+1}_{k-1}\right\}\\
&\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)f\,P^u_k=\frac{1}{2k+1}\left\{-\left(\frac{df}{dr}-\frac{k}{r}f\right)P^{u-1}_{k+1}+(k+u)(k+u-1)\left(\frac{df}{dr}+\frac{k+1}{r}f\right)P^{u-1}_{k-1}\right\}\\
&\frac{\partial}{\partial z}f\,P^u_k=\frac{1}{2k+1}\left\{\left(\frac{df}{dr}-\frac{k}{r}f\right)P^u_{k+1}+(k+u)(k-u)\left(\frac{df}{dr}+\frac{k+1}{r}f\right)P^u_{k-1}\right\}
\end{aligned}\right.$$

(7.2)

We may note that these formulae automatically look after "~~of~~ end effects"; thus, if we apply the first or third to P^k_k the factor (k−u) in the second term cuts out the terms in P^{k+1}_{k-1}, P^{k+1}_k, functions which does not exist.

Consider how these relations will work in (7,1). We try to get a solution in which at least one of the ψ's only involves a single spherical harmonic. Suppose that ψ_3 is proportional to P^u_k. Then the third equation of (7.1) tells us that terms from $\frac{\partial}{\partial z}\psi_1$ and $\left(\frac{\partial}{\partial x}\mp i\frac{\partial}{\partial y}\right)\psi_2$ must cancel out with this, and any other terms they give must cancel out together. It follows that ψ_1, ψ_2 must involve the same function of r and must either be of the forms P^u_{k+1}, P^{u+1}_{k+1}, or P^u_{k-1}, P^{u+1}_{k-1}. A similar argument then shows that ψ_4 must involve r in the same way as ψ_3 and must have spherical function P^{u+1}_k. We therefore take as trial solution

$$\left\{\begin{aligned}
&\psi_1=-ia_1F(r)\,P^u_{k+1},\ \Delta W_1 \qquad \psi_2=-ia_2F(r)\,P^{u+1}_{k+1}\ \Delta W_2\\
&\psi_3=a_3G(r)\,P^u_k,\ \Delta W_3 \qquad \psi_4=a_4G(r)\,P^{u+1}_k\ \Delta W_4
\end{aligned}\right.$$

$j = k+\frac{1}{2}\pm\frac{1}{2}$

$u = u$

(the factor −i is introduced in ψ_1, ψ_2 to make F real). We then find that the a's can be adjusted so that all four equations are satisfied. For example, the first equation gives

$$\frac{2\pi}{h}\left(\frac{W+eV}{c}+mc\right)a_1F\,P^u_{k+1}+\frac{a_4}{2k+1}\left\{-\left(\frac{dG}{dr}-\frac{k}{r}G\right)P^u_{k+1}+(k+u+1)(k+u)\times\right.$$
$$\left.\left(\frac{dG}{dr}+\frac{k+1}{r}G\right)P^u_{k-1}\right\}+\frac{a_3}{2k+1}\left\{\left(\frac{dG}{dr}-\frac{k}{r}G\right)P^u_{k+1}+(k+u)(k-u)\left(\frac{dG}{dr}+\frac{k+1}{r}G\right)P^u_{k-1}\right\}=0 \quad (7.3)$$

If, then, we take $a_4(k+u+1)+a_3(k-u)=0$, the terms in P^u_{k-1} cancel. In the second equation, the same ratio makes the coefficient of P^{u+1}_{k-1} disappear. From the other two equations we find similarly $a_1=a_2$; the ratio $a_1 : a_3$ is immaterial as it may be incorporated in F : G. We thus find as the complete solution

$$\left.\begin{aligned}
&\psi_1=-iF_k\,P^u_{k+1} \qquad \psi_2=-iF_k\,P^{u+1}_{k+1}\\
&\psi_3=(k+u+1)G_k\,P^u_k \qquad \psi_4=(-k+u)G_k\,P^{u+1}_k
\end{aligned}\right\}\quad(7.4)$$

where F_k, G_k satisfy the relations

$\frac{(j+1)}{r}+\frac{j+1}{r}\left(\frac{d}{dr}+\frac{k+2}{r}\right)\left(\frac{dG}{dr}-\frac{k}{r}\right)$

$\frac{d^2G}{dr^2}+\frac{k+2}{r}\left(\frac{k+2}{r}-\frac{k}{r}\right)\frac{dG}{dr}$

$+\frac{k}{r^2}G-\frac{k(k+2)}{r^2}G$

$1-k+2$ $=$ $-k(k+1)$

$$\left.\begin{aligned}\frac{2\pi}{h}\left(\frac{W+eV}{c}+mc\right)F+\frac{dG}{dr}-\frac{k}{r}G=0\\ -\frac{2\pi}{h}\left(\frac{W+eV}{c}-mc\right)G+\frac{dF}{dr}+\frac{k+2}{r}F=0\end{aligned}\right\}\qquad(7.5)$$

This solution we name $(k, j=k+\frac{1}{2}, m=u+\frac{1}{2})$.

A similar process gives a different solution if we make the first instead of the second terms in (7,3) cut out. We now have ψ_1 involving P^u_{k-1} instead of P^u_{k+1}. F and G must then satisfy

$$\left.\begin{aligned}\frac{2\pi}{h}\left(\frac{W+eV}{c}+mc\right)F+\frac{dG}{dr}+\frac{k+1}{r}G=0\\ -\frac{2\pi}{h}\left(\frac{W+eV}{c}-mc\right)G+\frac{dF}{dr}-\frac{k-1}{r}F=0\end{aligned}\right\}\qquad(7.6)$$

We can regard these equations as the same as (7.5) by changing k into $-k-1$; so we write the solution F_{-k-1}, G_{-k-1}. Then we have

$$\left.\begin{aligned}\psi_1=-i(k+u)F_{-k-1}P^u_{k-1},\quad&\psi_2=-i(-k+u+1)F_{-k-1}P^{u+1}_{k-1}\\ \psi_3=G_{-k-1}P^u_k\quad,&\quad\psi_4=G_{-k-1}P^{u+1}_k\end{aligned}\right\}\ (7,7)$$

$-k-1\rightarrow k-1$

This solution we name $(k, j=k-\frac{1}{2}, m=u+\frac{1}{2})$. We see in the subscripts of F and G the point of Dirac's method of allowing j to be negative. The equations (7.5) and (7.6) are substantially his equations at the foot of p. 622. We have not justified the use of k as a quantum number, and this cannot be done until we study intercombinations; but anticipating this ~~way~~ we may now count up the number of solutions and see that it is right for doublet spectra. In order to do this we must see what end cases are admissible. Take the first type of solution (7.4). If $u=k$, $\psi_4=0$ on account of the factor $(k+u)$, and so in substituting in (7.1) we shall not be led astray by applying (7.2) to the impermissible fn P^{k+1}_k. On the other hand, if we take $u=k+1$, we get a function ψ_2, but none of the others, and so evidently no solution. Similarly, at the other end we may take $u=-k-1$ (involving $\psi_3=0$) but not $u=-k-2$. In ~~the~~ all there will be $2k+2$ solutions. In the second type (7,7) we see in the same way that we may take u between $-k$ and $k-1$ inclusive and so get $2k$ solutions. In both cases there are therefore $2j+1$ solutions, as there should be. In the special case $k=0$ there are only two solutions of the

first type and none of the second.

We have thus found by trial a system of solution of our equations, and the important question arises as to whether it is a complete set. Can we simultaneously expand four arbitrary functions $\psi_1 \cdots \psi_4$ in terms of the solutions (7,4) and (7,7)? The full consideration would require a discussion of the radial functions including the quasi-hyperbolic case, which we ~~will~~ shall not attempt; for even without it we can see that we have only half as many proper functions as are required. Taking an arbitrary radius we may expand the four given functions as are required. Taking an arbitrary radius we may expand the four given functions in spherical harmonics over the ~~space~~ sphere. The kth harmonic will thus have $4(2k+1)$ coefficients, all arbitrary; whereas we have seen that there are only $2k+2+2k$ proper functions with ψ_3 and ψ_4 of order k. The incompleteness is evidently the same thing as was pointed out in §5. To get a complete ~~solution~~ set we must double the number of solutions by admitting negative values of the energy, and we have at present little idea of what this means.

8. We now discuss the radius the radial functions (7,5). In the case where the radial force is arbitrary we can proceed by approximation based on ~~by~~ the fact that F is much smaller than G. But the process would run very closely parallel to that of Dirac (p. 623) and we need not give it. We may only note that the null approximation gives Schrödinger's equation for G, and the next breaks it into two terms depending on parameters k and $-k-1$, in fact, the ordinary doublet spectrum.

We shall therefore proceed to find the accurate solution for the case of hydrogen. We take $V=e/r$ and it is convenient to introduce certain auxiliary symbols. Take

$$\frac{2\pi}{h}\left(mc+\frac{W}{c}\right)=A^2\quad,\quad\frac{2\pi}{h}\left(mc-\frac{W}{c}\right)=B^2\qquad(8,1)$$

with A and B both positive, and write as usual for the "fine structure constant

$$\gamma=\frac{2\pi e^2}{ch}\qquad(8,2)$$

$k=j$

and the equations (7.5) become

$$\left.\begin{aligned}\left(A^2+\frac{\gamma}{r}\right)F+\frac{dG}{dr}-\frac{k}{r}G=0\\ \left(B^2-\frac{\gamma}{r}\right)G+\frac{dF}{dr}+\frac{k+2}{r}F=0\end{aligned}\right\}\qquad(8,3)$$

We solve these in series of the form

$$\left.\begin{aligned} F &= e^{-\lambda r}\{a_0 r^\rho + a_1 r^{\rho-1} + a_2 r^{\rho-2} + \cdots\} \\ G &= e^{-\lambda r}\{b_0 r^\rho + b_1 r^{\rho-1} + b_2 r^{\rho-2} + \cdots\} \end{aligned}\right\} \quad (8.4)$$

Substitute and equate to zero the various terms. We have

$$A^2 a_0 - \lambda b_0 = 0 \qquad B^2 b_0 - \lambda a_0 = 0$$

$$A^2 a_1 + \gamma a_0 - \lambda b_1 + (\rho - k) b_0 = 0, \quad B^2 b_1 - \gamma b_0 - \lambda a_1 + (\rho + k + 2) a_0 = 0$$

$$\cdots$$

$$A^2 a_{s+1} + \gamma a_s - \lambda b_{s+1} + (\rho - k - s) b_s = 0, \quad B^2 b_{s+1} - \gamma b_s - \lambda a_{s+1} + (\rho + k + 2 - s) a_s = 0$$

The first pair determines $\lambda = AB$; we must take the positive solution to make F, G finite at infinity. We also have $b_0/a_0 = A/B$. Substituting in the next pair, we find that both a_1, b_1 can be eliminated simultaneously. We get the indical equation

$$\rho = -1 + \gamma \frac{A^2 - B^2}{2AB} \qquad (8,5)$$

A similar elimination of a_{s+1}, b_{s+1} can be carried out and we have

$$A a_s\left(\rho + k + 2 - s + \gamma \frac{B}{A}\right) + B b_s\left(\rho - k - s - \gamma\frac{A}{B}\right) = 0,$$

from which we may substitute

$$\left.\begin{aligned} b_s &= c_s\left[\gamma\frac{B}{A} + \rho + k + 2 - s\right] \\ a_s &= c_s \frac{B}{A}\left[\gamma\frac{A}{B} - \rho + k + s\right] \end{aligned}\right\} \qquad (8,6)$$

and can now form a difference equation for c_s. This reduces to

$$AB\,(2s+2)c_{s+1} = -c_s\{(s-\rho-1)^2 - [(k+1)^2 - \gamma^2]\}.$$

Writing $\sqrt{(k+1)^2 - \gamma^2} = k'$ (suppose positive) we have

$$2AB\,(s+1)\,c_{s+1} = -c_s\,(\rho + 1 - s - k')(\rho + 1 - s + k')$$

and so

$$c_s = (-)^s \frac{(\rho - k' + 1)(\rho - k')\cdots(\rho - k' - s + 2)(\rho + k' + 1)(\rho + k')\cdots(\rho + k' - s + 2)}{2^s \cdot s!\,(AB)^s} \quad (8.7)$$

The series for F and G are composed of terms of this type each multiplied by a factor given by (8,6). If the solution is to be finite throughout space it is necessary that these series should terminate for some value of s such that $\rho - s > 0$. It is therefore necessary that $\rho = k' + n' - 1$, where

n' is zero or a positive integer. This condition determines the energy levels. For we have

$$k' + n' = \gamma\frac{A^2 - B^2}{2AB} = \gamma\frac{W}{\sqrt{\{m^2c^4 - W^2\}}},$$

and so

$$W = mc^2\left\{1 + \frac{\gamma^2}{(k'+n')^2}\right\}^{-\frac{1}{2}} \qquad (8,8)$$

This is exactly the original Sommerfeld expression for the energy levels of hydrogen. The only difference is that our k may take the value zero, so that the formula now involves $\sqrt{\{(k+1)^2 - \gamma^2\}}$.

The process of solution has at no stage made use of the fact that k is positive, and we conclude that the same solution will hold for (7, 6) provided that we write $-k-1$ for k. In this case then $k' = \sqrt{k^2 - \gamma^2}$, and we see how the two k levels split. If we compare j's instead of k's we must take $k+1$ for the second solution, and the levels fall exactly together. In the case $k=0$ there is no second solution and no corresponding radial function for $k = -1$, as (8,7) will not then factorise into real factorials, so that the series cannot terminate.

We may now express the properfunctions in terms of the quantum numbers. We make use of the auxiliary quantities.

$$k' = \sqrt{\{k+1\}^2 - \gamma^2}, \qquad N = \sqrt{\{(k'+n')^2 + \gamma^2\}}$$

$k=0$

and

$$\frac{2\pi mc\gamma}{h} = \frac{4\pi^2 m e^2}{h^2} = \frac{1}{a}.$$

Then a is the "radius of the first hydrogen orbit" and N is approximately the total quantum number, counting 1, 2, 3 ⋯ for the hydrogen levels. We then have

$$W = mc^2(k'+n')/N$$

$$G_k = e^{-r/aN}\left\{r^{k'+n'-1}(N+\overset{k'}{k}+1) - r^{k'+n'-2}\,aN(N+k)\frac{n'(n'+2k')}{2} + r^{k'+n'-3}a^2N^2(N+k-1)\frac{n'(n'-1)(n'+2k')(n'+2k'-1)}{2\cdot 4} - \cdots\right\}$$

$$F_k = \frac{\gamma}{N+k'+n'}e^{-r/aN}\left\{r^{k'+n'-1}(N+k+1) - r^{k'+n'-2}\,aN(N+k+2)\frac{n'(n'+2k')}{2} + r^{k'+n'-3}a^2N^2(N+k+3)\frac{n'(n'-1)(n'+2k')(n'+2k'-1)}{2\cdot 4} - \cdots\right\}$$

$$AB = \frac{2\pi}{h}\cdot\sqrt{m^2c^2 - \frac{W^2}{c^2}}$$

$$= \frac{2\pi mc}{h}\cdot\sqrt{1 - \frac{(k'+n')^2}{N^2}} = \frac{2\pi mc}{h}\,\frac{\gamma}{N} = \frac{1}{aN}$$

$$F_k = \frac{r}{N+k'+n'} e^{-r/aN} \Big\{ r^{k'+n'-1}(N+k+1) - r^{k'+n'-2} aN(N+k+2)\frac{n'(n'+2k')}{2}$$
$$+ r^{k'+n'-3} a^2N^2(N+k+3)\frac{n'(n'-1)(n'+2k')(n'+2k'-1)}{2\cdot 4} - \cdots \Big\}.$$

We may observe that if we approximate by neglecting γ, we find that G_k and G_{-k-1} are respectively $(N+k+1)$ and $(N-k)$ times the ordinary radial function of Schrödinger.

The solution reveals a small blemish in the equations, for we have to admit the existence of proper functions which become infinite. The last term in the series for G has power $r^{k'-1}$, and if $k=0$, k' is very slightly less than 1, so that there will be a term with a small negative power of r. Of course all integrals connected with the spectrum are amply convergent. We do not perhaps known enough about the essential rules for proper functions to pay much attention to this defect. Moreover, it may well be that it would disappear if could solve the problem of two bodies properly instead of treating the nucleus as an abstract centre of force.

As an example of these apparently complicated functions, we may set down the solutions corresponding to $N=1$ (exactly), the lowest hydrogen state. We shall replace the spherical harmonics symbols by the corresponding solid harmonics. Then the two solutions are $(k=0, n'=0, N=1, u=0, \text{and} -1)$

$$W = mc^2\sqrt{(1-\gamma^2)} \qquad \rho = \sqrt{(1-\gamma^2)} - 1$$

$$\psi_1 = -\frac{i\gamma}{1+\sqrt{(1-\gamma^2)}} 2r^{\rho-1}e^{-r/a}$$

$$\psi_2 = -\frac{i\gamma}{1+\sqrt{(1-\gamma^2)}}(x+iy)r^{\rho-1}e^{-r/a},$$

$$\psi_3 = r^\rho e^{-r/a}, \qquad \psi_4 = 0,$$

and
$$\psi_1 = \frac{i\gamma}{1+\sqrt{1-\gamma^2}}(x-iy)r^{\rho-1}e^{-r/a}$$

$$\psi_2 = -\frac{i\gamma}{1+\sqrt{(1-\gamma^2)}} 2r^{\rho-1}e^{-r/a},$$

$$\psi_3 = 0 \qquad \psi_4 = -r^\rho e^{-r/a}$$

$Y_3 = -i\frac{\gamma}{1+\sqrt{1-\gamma^2}}\frac{z}{r}r^\rho e^{-r/a}$

$Y_1 = \qquad \frac{x}{r}$

$Y_2 = \qquad \frac{y}{r}$

$X_0 = r^\rho e^{-r/a}$

~~Y_4~~ $X_1 = X_2 = X_3 = 0$!!

$$\boxed{Y_r = \frac{-i\gamma}{1+\sqrt{1-\gamma^2}} r^\rho e^{-r/a}}$$

This will suffice as an illustration of the accurate solution.

9. We ~~may~~ now consider the rules of combination. The emission can be calculated from (3,2) by setting down the values of V and A at a distant point. They depend on the retarded potentials and the work follows that of Klein very closely, so that we need not give details. We omitt discussion of the very weak radiations of quadrupole and higher types. In calculating the chief radiation we shall have contributions from V and also from A. Taking the transition $p \to q$ we write $W_p - W_q = h\nu$ and find

$$V(x',y',z') = \frac{e^{i2\pi\nu(t-r'/c)}}{r'}\iiint \rho_{pq}\frac{(\mathbf{r}.\mathbf{r}')}{r'} i\frac{2\pi\nu}{c}\,dx\,dy\,dz,$$

where $\mathbf{r}, \mathbf{r}'$ are the vectors from the origin to x, y, z and the distant point of observation x′, y′, z′ respectively, and r′ is the absolute value of $\mathbf{r}'$.
Similarly

$$A_1(x',y',z') = \frac{e^{i2\pi\nu(t-r'/c)}}{r'}\iiint \frac{(j_1)_{pq}}{c}\,dx\,dy\,dz.$$

V involves the electric moment, and it should be noticed that in spite of its different appearance, A does so too. For

$$\iiint j_1\,dx\,dy\,dz = \iiint -x\Big(\frac{\partial j_1}{\partial x} + \frac{\partial j_2}{\partial y} + \frac{\partial j_3}{\partial z}\Big)dx\,dy\,dz.$$

~~V involves the electric moment, and it should be noticed that, in spite of its different app~~ since on partial integration the last two terms vanish. Hence by (3,5)

$$\iiint j_1\,dx\,dy\,dz = \iiint x\frac{\partial\rho}{\partial t}\,dx\,dy\,dz = i2\pi\nu\iiint x\rho\,dx\,dy\,dz$$

Thus it will suffices to discuss the electric moment.
When our proper functions are written as the tesseral spherical harmonics, the three appropriate types of moment are $x+iy$, z, $x-iy$. For determining what combinations occur we require the following easily proving relations:—

$$\iint P_k^u P_{k-1}^{u-1*}\sin\theta\, e^{-i\phi}\sin\theta\, d\theta\, d\phi = -\iint P_k^u P_{k-1}^{u+1*}\sin\theta\, e^{i\phi}\sin\theta\, d\theta\, d\phi$$
$$= \iint P_k^u P_{k-1}^{u*}\underset{\sin\theta}{\cos\theta}\, d\theta\, d\phi = \frac{4\pi}{(2k+1)(2k-1)}(k+u)!(k-u)! \qquad (9.1)$$

For all other products of these types the integral vanishes. We may here note also the normalising relation.

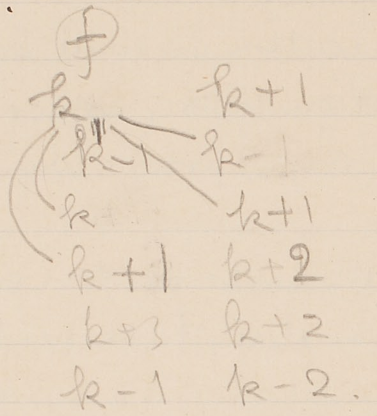
$H\Psi+V\Psi=E\psi_1$

$\psi_1\ \psi_2\ \psi_3\ \psi_4$

$$\iint P_k^u P_k^{u*}\sin\theta\,d\theta\,d\phi=\frac{4\pi}{2k+1}(k+u+1)!(k-u)! \qquad (9.2)$$

With the help of (9.1) we can see what combinations might occur. It will evidently suffice to treat of only one type of polarisation and we shall take that corresponding to z. The following scheme then shows the only solutions which might, according to (9.1), combine with the first. We only mark the harmonic coefficients.

k	j	ψ_1	ψ_2	ψ_3	ψ_4
k	$k+\frac{1}{2}$	P^u_{k+1}	P^{u+1}_{k+1}	P^u_k	P^{u+1}_k
$k-1$	$k-\frac{1}{2}$	P^u_k	P^{u+1}_k	P^u_{k-1}	P^{u+1}_{k-1}
$k+1$	$k+\frac{1}{2}$	P^u_k	P^{u+1}_k	P^u_{k+1}	P^{u+1}_{k+1}
$k+1$	$k+\frac{3}{2}$	P^u_{k+2}	P^{u+1}_{k+2}	P^u_{k+1}	P^{u+1}_{k+1}
$k+3$	$k+\frac{5}{2}$	P^u_{k+2}	P^{u+1}_{k+2}	P^u_{k+3}	P^{u+1}_{k+3}
$k-1$	$k-\frac{3}{2}$	P^u_{k-2}	P^{u+1}_{k-2}	P^u_{k-1}	P^{u+1}_{k-1}



When we examine these, actually putting in the coefficients, we find that the last two vanish identically. This verifies the j rules, that only $j\to j\pm1$ or $j\to j$ are possible combinations,+ a similar results follow, of course, if we examine the other polarisations or the combinations of the second type of solution. We shall not give the numerical values here as these are well known, but one more point deserves remark. The radial integration are, speaking accurately, different for the three lines arising from a given $k\to(k-1)$. They are, in fact, $\int(F_kF_{k\pm1}+G_kG_{k-1})r^3dr$ $\int(F_{-k-1}F_{-k}+G_{-k-1}G_k)r^3dr$ and

$$\int(F_{-k-1}F_{k-1}+G_{-k-1}G_{k-1})r^3dr.$$

But to a first approximation we saw that F could be neglected beside G, while $G_{-k-1}=G_k$, so that to this approximation the three radial integrations are the same. This explains why the intensities bear simple numerical ratios to one another in doublet spectra.

10. When a uniform magnetic field is imposed on a doublet atom it is not possible to get an accurate solution, and we have to fall back on the method of

+ Of course the k rule is required as well, for there are levels k, $k-\frac{1}{2}$ and $k+2$, $k+\frac{3}{2}$ which do not combine with k, $k+\frac{1}{2}$. Dirac's use of negative j does make the statements simpler.

perturbations. The simplest way of working out the Zeeman effect is to use the approximation of §4, which reduces it to the work done in my earlier paper. That this is a sufficient approximation may be seen from the fact that it gives the doublet fine structure and treats the magnetic structure as of the same order of magnitude, which is just the degree of accuracy required to explain the observed effects. But a direct attack, starting from the accurate solution of (7.1) and superposing on it the magnetic field is also interesting; it throws the solution into rather a different form because the levels of the fine structure are already separated, whereas in the earlier process they were attributed to a perturbation acting together with the magnetic field.

We must first see how the methods of perturbations will go. The solutions of (2,2) when p_0 is replaced by W so that they give the levels, obey an orthogonal relation, as is easily proved directly; thus for any two solutions p, q either

$$W_p-W_q=0 \quad \text{or} \quad \iiint\sum_{\lambda=1}^{4}\psi_\lambda^p\psi_\lambda^{q*}\,dx\,dy\,dz=0. \qquad (10,1)$$

In the case of degeneracy, where $W_p=W_q$, the partly arbitrary ψ_λ^q can be chosen so that (10,1) will still be true. This has already been done in our case. Let us suppose that on account of small changes in V and A the four ~~quantities~~ equations (2,2) are affected by small extra terms which we may write as $P_1\psi$ .. $P_4\psi$, each symbol signifying that any of the ψ's may enter into each equation. We require a solution near W_p and assume that it is of the form

$$\psi_\lambda=\sum a_s\psi_\lambda^s,$$

where the a_s's are small for all cases where W_s is not nearly equal to W_p. Notice the slight difference that we must make from the theory of perturbations of a degenerate system, on account of the fact that we have to allow for the interaction of solutions that initially are not quite coincident. Substituting in (2,2) we have

$$\sum_s\frac{1}{c}(W-W_s)a_s\psi_1^s+\sum_s a_sP_1\psi^s=0, \quad \text{etc.}$$

We multiply these by ψ_1^{q*}, etc., add the four equations together and integrate over space. We thus find

Now we have seen that in all these the terms in F_k, F_{-k-1} are small compared to those in G. Further~~f~~ more, we know that to the same approximation G_{-k-1} is proportional to G_k, and may be taken equal to it by a suitable definition. Hence all the radial integrations are the same, and they may therefore be omitted. The equation (10,2) thus becomes

$$\frac{a}{c}(W-W_p) - \frac{eH}{2c}\frac{h}{4\pi mc}\frac{2k+2}{2k+1}(2u+1)a - \frac{eH}{2c}\frac{h}{4\pi mc}\frac{2b}{2k+1} = 0$$

$$\frac{b}{c}(W-W_q) - \frac{eH}{2c}\frac{h}{4\pi mc}\frac{2k}{2k+1}(2u+1)b - \frac{eH}{2c}\frac{h}{4\pi mc}\frac{(k+u+1)(k-u)}{2k+1}2a = 0$$

To reduce this to the familiar form we write W_0 as the mean centre of the two lines, so that

$$W_p = W_0 + k\rho, \qquad W_q = W_0 - (k+1)\rho,$$

and $W = W_0 + \overline{W}$. Also take $\omega = \frac{eH}{2mc}\frac{h}{2\pi}$. Then we have

$$a\left\{\overline{W} - k\rho - \omega\frac{k+1}{2k+1}(2u+1)\right\} - b\omega\frac{1}{2k+1} = 0,$$

$$-a\omega\frac{(k+u+1)(k-u)}{2k+1} + b\left\{\overline{W} + (k+1)\rho - \omega\frac{k}{2k+1}(2u+1)\right\} = 0,$$

and from this we derive

$$\overline{W}^2 + \rho\overline{W} - k(k+1)\rho^2 + \omega(2u+1)(\overline{W}+\rho) + \omega^2 u(u+1) = 0,$$

which is the standard equation for the Zeeman effect in doublets. It will be seen that the algebra is a little more complicated than that which comes from forming the approximate equations in §3, §4 as was done before. We shall not work out intensities as nothing new would be found.

Summary

$$\frac{1}{c}(W-W_q)a_q\int\sum_{\lambda=1}^{4}\psi_\lambda^q\psi_\lambda^{q*}\,dx\,dy\,dz+\sum_s a_s\int\sum_{\lambda=1}^{4}\psi_\lambda^{q*}P_\lambda\psi^s\,dx\,dy\,dz=0$$

If there are a number of states q near p, even though not coincident with it, the associated a_q's need not be small and must be retained. We then form the determinant in the usual way, giving an algebraic equation for W, and afterwards we can determine the various ratios for a_q.

For a magnetic field along z we take $A_1=-\frac{1}{2}Hy$, $A_2=\frac{1}{2}Hx$ and so must add on to (2,2) terms

$$P_1\psi=-i\frac{eH}{2c}(x-iy)\psi_4,\qquad P_2\psi=i\frac{eH}{2c}(x+iy)\psi_3,$$
$$P_3\psi=-i\frac{eH}{2c}(x-iy)\psi_2,\qquad P_4\psi=i\frac{eH}{2c}(x+iy)\psi_1.$$

We hence get

$$\frac{1}{c}(W-W_q)a_q\int\sum_{\lambda=1}^{4}\psi_\lambda^q\psi_\lambda^{q*}\,dx\,dy\,dz+\frac{eH}{2c}\sum_s a_s[q;s]=0 \qquad (10,2)$$

where

$$[q;s]=\int\{-i\psi_1^{q*}(x-iy)\psi_4^s+i\psi_2^{q*}(x+iy)\psi_3^s-i\psi_3^{q*}(x-iy)\psi_2^s+i\psi_4^{q*}(x+iy)\psi_1^s\}\,dx\,dy\,dz \qquad (10.3)$$

This expression determines whether two levels q, s can interfere with one another in producing the Zeeman effect. If we apply the formulae of (9.1) to any of the solutions (7,4) or (7,7), we at once see that unless the number u is the same for both q and s, the integral must vanish. Hence only levels with the same quantum number $m\ (=u+\frac{1}{2})$ can interfere. Further, we can build a table, of the same kind as was made in §9 for combinations, to show what possible k, j values might give non-vanishing integrals, and just as there, we find that some of the possibilityies disappear on closer examination. We are left to consider interferences of $(k, k+\frac{1}{2})$ and $(k, k-\frac{1}{2})$ with themselves and one another and also of $(k, k+\frac{1}{2})$ with $(k+2, k+\frac{3}{2})$ (the case $(k, k-\frac{1}{2})$ with $(k-\frac{3}{2}, k-\frac{5}{2})$, which also occurs, is essentially the same as this last). The first set are what we are familiar with, but the last is rather unexpected. We shall prove that it is small of an order to be neglected, because it corresponds to the terms we should have by a second approximation.

Substituting out of (7,4), (7,7) and applying (9.1) we find

$$[k, k+\tfrac{1}{2};\ k+2, k+\tfrac{3}{2}]=\frac{4\pi}{2k+3}(k+u+2)!\,(k-u+1)!\int(F_kG_{k-3}+G_kF_{k-3})\times r^3\,dr.$$

In order to reduce the radial integration, we proceed as follows. Take the equations (7,5) and a similar pair, in which k is replaced by another number l. Multiply the first of (7.5) by G_l, the second by F_l and add. This removes the terms in W and eV. Then interchange k and l and add the two expressions together. The result is

$$\frac{2\pi}{h}\cdot 2mc\,(F_kG_l+F_lG_k)+\frac{d}{dr}(F_kF_l+G_kG_l)-\frac{k+l}{r}G_kG_l+\frac{k+l+4}{r}F_kF_l=0.$$

Hence making a partial integration we find

$$\int(F_kG_l+F_lG_k)\,r^3dr=\frac{h}{4\pi mc}\int[G_kG_l(k+l+3)-F_kF_l(k+l+1)]\,r^2dr \qquad (10,4)$$

For the case $l=-k-3$, the term in G therefore vanishes. We know that F is smaller than G in a ratio of order $\gamma:1$; hence this integral bears to one where the coefficient of G does not vanish a ratio $\gamma^2:1$, which is the order of a term of the second not the first approximation.

We are thus left with the result that only the two k levels interfere. Let us call $(k, k+\frac{1}{2})$ p and $(k, k-\frac{1}{2})$ q. Then our solution may be written as

$$\psi_\lambda=a\psi_\lambda^p+b\psi_\lambda^q,$$

and we find by the application of (9,1) and (10,4) that

$$[p;q]=-4\pi\frac{2k+2}{2k+1}(2u+1)(k+u+1)!\,(k-u)!\,\frac{h}{4\pi mc}\int\left(G_k^2-\frac{2k+1}{2k+3}F_k^2\right)r^2dr,$$
$$[q;q]=-4\pi\frac{2k}{2k+1}(2u+1)(k+u)!\,(k-u-1)!\,\frac{h}{4\pi mc}\int\left(G_{-k-1}-\frac{2k+1}{2k-1}F_{-k-1}\right)\times r^2dr,$$
$$[p;q]=[q;p]=-\frac{4\pi}{2k+1}\,2(k+u+1)!\,(k-u)!\,\frac{h}{2\pi mc}\int G_kG_{-k-1}\,r^2dr.$$

We also require the normalisations. Using (9.2) these are

$$\int\sum_\lambda|\psi_\lambda^p|^2\,dx\,dy\,dz=4\pi(k+u+1)!\,(k-u)!\int(F_k^2+G_k^2)\,r^2dr=1.$$
$$\int\sum_\lambda|\psi_\lambda^q|^2\,dx\,dy\,dz=4\pi(k+u)!\,(k-u-1)!\int(F_{-k-1}^2+G_{-k-1}^2)\,r^2dr.$$