PROC. ROY. SOC. (LONDON) A 117 (1928) 610 – 624.

The Quantum Theory of the Electron.

By P. A. M. DIRAC, St. John's College, Cambridge.

(Communicated by R. H. Fowler, F.R.S.—Received January 2, 1928.)

The new quantum mechanics, when applied to the problem of the structure of the atom with point-charge electrons, does not give 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, Goudsmit 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 order 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, alternatetively, with the general transformation theory of quantum mechanics. It appears that the simplest Hamiltonian for a point-charge electron 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, 'Z. f. Physik,' vol. 43, p. 601 (1927).
† Darwin, 'Roy. Soc. Proc.,' A, vol. 116, p. 227 (1927).

§ 1. *Previous Relativity Treatments.*

The relativity Hamiltonian according to the classical theory for a point electron moving in an arbitrary electro-magnetic field with scalar potential A_0 and vector potential $\mathbf{A}$ is

$$F \equiv \left(\frac{W}{c} + \frac{e}{c} A_0\right)^2 + \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2,$$

where $\mathbf{p}$ is the momentum vector. It has been suggested by Gordon* that the operator of the wave equation of the quantum theory should be obtained from this F by the same procedure as in non-relativity theory, namely, by putting

$$W = ih\frac{\partial}{\partial t},$$

$$p_r = -ih\frac{\partial}{\partial x_r}, \qquad r = 1, 2, 3,$$

in it. This gives the wave equation

$$F\psi \equiv \left[\left(ih\frac{\partial}{c\,\partial t} + \frac{e}{c}A_0\right)^2 + \Sigma_r\left(-ih\frac{\partial}{\partial x_r} + \frac{e}{c}A_r\right)^2 + m^2c^2\right]\psi = 0, \qquad (1)$$

the wave function ψ being a function of x_1, x_2, x_3, t. This gives rise to two difficulties.

The first is in connection with the physical interpretation of ψ. Gordon, and also independently Klein,† from considerations of the conservation theorems, make the assumption that if ψ_m, ψ_n are two solutions

$$\rho_{mn} = -\frac{e}{2mc^2}\left\{ih\left(\psi_m\frac{\partial\bar{\psi}_n}{\partial t} - \bar{\psi}_n\frac{\partial\psi_m}{\partial t}\right) + 2eA_0\psi_m\bar{\psi}_n\right\}$$

and

$$\mathbf{I}_{mn} = -\frac{e}{2m}\left\{-ih\left(\psi_m \operatorname{grad}\bar{\psi}_n - \bar{\psi}_n \operatorname{grad}\psi_m\right) + 2\frac{e}{c}\mathbf{A}_m\psi_m\bar{\psi}_n\right\}$$

are to be interpreted as the charge and current associated with the transition $m \rightarrow n$. This appears to be satisfactory so far as emission and absorption 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

* Gordon, 'Z. f. Physik,' vol. 40, p. 117 (1926).

† Klein, 'Z. f. Physik,' vol. 41, p. 407 (1927).

‡ Jordan, 'Z. f. Physik,' vol. 40, p. 809 (1927); Dirac, 'Roy. Soc. Proc.,' A, vol. 113, p. 621 (1927).

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

$$(\mathrm{H} - \mathrm{W})\,\psi = 0, \qquad (2)$$

i.e., being linear in W or $\partial/\partial t$, 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 equation (1), one gets

$$\left[\left(-\frac{\mathrm{W}}{c} + \frac{e}{c}\mathrm{A}_0\right)^2 + \left(-\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2\right]\psi = 0,$$

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 definiteness 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 transition would appear experimentally as the electron suddenly changing its charge 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

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

if one puts

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

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

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

where for the present all that is known about the dynamical variables or operators α_1, α_2, α_3, β is 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 α_1, α_2, α_3, β are independent of t, x_1, x_2, x_3, *i.e.*, that they commute 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 α_1, α_2, α_3, β 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

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

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 put $\beta = \alpha_4 mc$, these conditions become

$$\alpha_\mu^2 = 1 \qquad \alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0 \;(\mu \neq \nu) \qquad \mu, \nu = 1, 2, 3, 4. \tag{6}$$

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

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

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 :—

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

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

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)

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

* Pauli, *loc. cit.*

If we now take

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

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

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

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

The wave equation (4) now takes the form

$$[p_0 + \rho_1(\boldsymbol{\sigma}, \mathbf{p}) + \rho_3 mc]\,\psi = 0, \tag{9}$$

where $\boldsymbol{\sigma}$ 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

$$p_0 = ip_4,$$

$$\rho_3 = \gamma_4, \qquad \rho_2\sigma_r = \gamma_r, \qquad r = 1, 2, 3, \tag{10}$$

we have

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

The p_μ transform under a Lorentz transformation according to the law

$$p_\mu' = \Sigma_\nu a_{\mu\nu} p_\nu,$$

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

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

The wave equation therefore transforms into

$$[i\Sigma\gamma_\mu' p_\mu' + mc]\,\psi = 0, \tag{12}$$

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

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

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 corresponding 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 characteristic values 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 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

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

where the sixteen coefficients c, c_r, c_r', c_{rs} are c-numbers. By expressing σ_1' in this way, we see, 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

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

* We say that a anticommutes with b when $ab = -ba$.

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 condition $\sigma_1'^2 = 1$ shows that $a_{12}a_{21} = 1$, $a_{34}a_{43} = 1$. 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 $(a_{43}/a_{34})^{\frac{1}{2}}$, and 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 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 now apply the canonical transformation: first and second rows to be multiplied by $e^{i\theta}$ and first and second columns to be divided by the same expression, ρ_1' will be brought into the form ρ_1, and σ_1, σ_3, ρ_3 will not be altered. ρ_2' and σ_2' must now 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 ρ's and σ's. The new wave equation (12) can in this way be brought back into the form of the original wave equation (11) or (9), so that the results that follow from this original wave equation must be independent of the frame of reference used.

§ 4. *The Hamiltonian for an Arbitrary Field.*

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

$$\left[p_0 + \frac{e}{c}A_0 + \rho_1\left(\boldsymbol{\sigma}, \mathbf{p} + \frac{e}{c}\mathbf{A}\right) + \rho_3 mc\right]\psi = 0. \tag{14}$$

This wave equation 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 many 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 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

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

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

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

Thus (15) becomes

$$0 = \Big[-(p_0 + e'\mathrm{A}_0)^2 + (\mathbf{p} + e'\mathbf{A})^2 + m^2c^2 + e'h(\boldsymbol{\sigma}, \operatorname{curl} \mathbf{A}) - ie'h\rho_1\Big(\boldsymbol{\sigma}, \operatorname{grad} \mathrm{A}_0 + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\Big)\Big]\psi$$

$$= [-(p_0 + e'\mathrm{A}_0)^2 + (\mathbf{p} + e'\mathbf{A})^2 + m^2c^2 + e'h(\boldsymbol{\sigma}, \mathbf{H}) + ie'h\rho_1(\boldsymbol{\sigma}, \mathbf{E})]\psi,$$

where **E** and **H** are the electric and magnetic vectors 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 behave as though it has a magnetic moment $eh/2mc \,.\, \boldsymbol{\sigma}$ and an electric moment $ieh/2mc \,.\, \rho_1\boldsymbol{\sigma}$. 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 is 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.

§ 5. *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 $e'\mathrm{A}_0 = \mathrm{V}(r)$, an arbitrary function of the radius r, so that the Hamiltonian in (14) becomes

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

We shall determine the periodic solutions of the wave equation $\mathrm{F}\psi = 0$, 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 **m** is defined by

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

and satisfies the following " Vertauschungs " relations

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

together with similar relations obtained by permuting the suffixes. Also $\mathbf{m}$ commutes with r, and with p_r, the momentum canonically conjugate to r.

We have

$$m_1 \mathrm{F} - \mathrm{F} m_1 = \rho_1 \{ m_1 (\boldsymbol{\sigma}, \mathbf{p}) - (\boldsymbol{\sigma}, \mathbf{p}) m_1 \}$$
$$= \rho_1 (\boldsymbol{\sigma}, m_1 \mathbf{p} - \mathbf{p} m_1)$$
$$= ih\rho_1 (\sigma_2 p_3 - \sigma_3 p_2),$$

and so

$$\mathbf{m}\mathrm{F} - \mathrm{F}\mathbf{m} = ih\rho_1 \, \boldsymbol{\sigma} \times \mathbf{p}. \tag{18}$$

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

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

with the help of (8), and so

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

Hence

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

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

$$\mathbf{M} \times \mathbf{M} = ih\mathbf{M} \quad \text{and} \quad \mathrm{M}^2 \mathrm{M}_3 = \mathrm{M}_3 \mathrm{M}^2.$$

M_3 will be an action variable of the system. Since the characteristic values of m_3 must be integral multiples of h in 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

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

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, and defines the azimuthal 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.

* See 'Roy. Soc. Proc.,' A, vol. 111, p. 281 (1926).

§ 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 $\mathbf{B} = \mathbf{C} = \mathbf{m}$. This gives

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

Hence

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

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

$$\begin{aligned}(\boldsymbol{\sigma}, \mathbf{m})(\boldsymbol{\sigma}, \mathbf{p}) + (\boldsymbol{\sigma}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{m}) &= i\Sigma\sigma_1(m_2p_3 - m_3p_2 + p_2m_3 - p_3m_2) \\ &= i\Sigma\sigma_1 \,.\, 2ihp_1 = -2h(\boldsymbol{\sigma}, \mathbf{p}),\end{aligned}$$

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, $\rho_1(\boldsymbol{\sigma}, \mathbf{p})$, 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^2h^2. We therefore take

$$jh = \rho_3\{(\boldsymbol{\sigma}, \mathbf{m}) + h\}. \tag{21}$$

We have, by a further application of (16)

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

Now a permissible definition of p_r is

$$(\mathbf{x}, \mathbf{p}) = rp_r + ih,$$

and from (21)

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

Hence

$$(\boldsymbol{\sigma}, \mathbf{x})(\boldsymbol{\sigma}, \mathbf{p}) = rp_r + i\rho_3 jh. \tag{22}$$

Introduce the quantity ε defined by

$$r\varepsilon = \rho_1 (\boldsymbol{\sigma}, \mathbf{x}). \tag{23}$$

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

$$r^2\varepsilon^2 = [\rho_1 (\boldsymbol{\sigma}, \mathbf{x})]^2 = (\boldsymbol{\sigma}, \mathbf{x})^2 = \mathbf{x}^2 = r^2$$

or

$$\varepsilon^2 = 1.$$

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

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

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

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

or

$$\rho_1 (\boldsymbol{\sigma}, \mathbf{p}) = \varepsilon p_r + i\varepsilon p_3 jh/r.$$

Thus

$$\mathrm{F} = p_0 + \mathrm{V} + \varepsilon p_r + i\varepsilon\rho_3 jh/r + \rho_3 mc. \tag{24}$$

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 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

$$\mathrm{F}\psi \equiv [p_0 + \mathrm{V} + \rho_2 p_r - \rho_1 jh/r + \rho_3 mc]\,\psi = 0.$$

If we write this equation 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

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

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

The second and fourth components give just a repetition of these two equations. We shall now eliminate ψ_α. If we write $h\mathrm{B}$ for $p_0 + \mathrm{V} + mc$, the first equation becomes

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

which gives on differentiating

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

This reduces to

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

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

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

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 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

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

which gives

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

The correction is now, to the first order of accuracy

$$\frac{1}{\mathrm{B}h}\left(\frac{j}{r}\frac{\partial \mathrm{V}}{\partial r} - \tfrac{1}{2}\frac{\partial^2 \mathrm{V}}{\partial r^2}\right),$$

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

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

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

$$\frac{e^2}{2mc^2r^3}\,j \tag{28$'$}$$

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

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

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

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

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

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

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

or

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

hence

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

The correction thus 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 first approximation, lead to the same energy levels as those obtained by Darwin, which are in agreement with experiment.

Einige Bemerkungen zur Diracschen Theorie des relativistischen Drehelektrons.

Von J. v. Neumann in Berlin.

(Eingegangen am 15. März 1928.)

Einige Eigenschaften des Diracschen Drehelektrons werden näher analysiert, und zwar die Natur der monochromatischen de Broglie-Wellen, die Transformationseigenschaften der vier ψ-Komponenten, sowie der Energie-Strom-Vektor (dessen Zeitkomponente die Wahrscheinlichkeit ist*).

Einleitung.

I. Herr P. A. M. Dirac hat in einer kürzlich erschienenen Arbeit** eine neue Behandlungsweise des quantenmechanischen Einkörperproblems vorgeschlagen, bei der gewisse Mängel der bisherigen relativistischen Einkörper-Gleichung*** auf einfache und befriedigende Weise behoben wurden. Sein Ansatz ermöglichte aber nicht nur die Behebung der genannten (relativistischen) Mängel, er ergab vielmehr auch — wie er l. c. zeigte — ganz von selbst die bekannten Spin-Eigenschaften des Elektrons: nämlich sein mechanisches Drehmoment $\frac{h}{4\pi}$, sowie seine magnetischen Momente $\frac{he}{8\pi mc}$, $\frac{he}{4\pi mc}$ im „inneren" bzw. „äußeren" Felde****.

Dieser überwältigende Erfolg der Diracschen Theorie läßt wohl kaum Zweifel in der Beziehung, daß mit ihr wesentliche Merkmale des relativistischen Verhaltens von „Massenpunkten" erfaßt sind* (wenn auch noch gewisse, von Dirac mit voller Schärfe hervorgehobene und gleich zu nennende Schwierigkeiten zunächst fortbestehen), und es ist daher wohl angezeigt, sich mit den Konsequenzen derselben im Detail auseinanderzusetzen. Denn die heutige „Transformationstheorie" der Quantenmechanik ist ihrem Wesen nach (im besonderen im ihr zugrunde liegenden Wahrscheinlichkeitsbegriff) eine unrelativistische, auf der Idee der objektiven Gleichzeitigkeit beruhende**: es ist daher notwendig festzustellen, wie sie sich unter dem Einfluß eines durch und durch relativistischen Modells vom „Massenpunkt" verwandelt.

Außerdem ist in die Diracsche Theorie ein unangenehmes Erbstück aus der früheren (vgl. Anm.***, S. 868) relativistischen Gleichung mitgekommen: nämlich die Unbestimmtheit des Vorzeichens der Ladung des Elektrons, bzw. der Richtung des Zeitablaufs (diese verursacht es, daß Diracs Wellenfunktion nicht von zwei Wellenfunktionen im Schrödingerschen Sinne — wie es das Auftreten des Spins erfordern würde — gebildet wird, sondern von vier)***. Auch die Rolle dieser Komplikation soll uns beschäftigen, wenn wir auch keineswegs in der Lage sind, sie erschöpfend zu behandeln.

Zum Schlusse will ich noch darauf hinweisen, daß ich beim Entstehen dieser Arbeit wesentliche Anregungen in Gesprächen mit Herrn P. Jordan und E. Wigner in Göttingen empfangen habe, denen ich hier wärmstens danken möchte.

Allgemeines.

Monochromatisch-ebene Wellen in der Diracschen Theorie.

II. Diracs Theorie beruht bekanntlich auf dem folgenden Prinzip: Die alte relativistische Gleichung (ohne Kraftfelder!)

$$\left(p^2 + q^2 + r^2 - \frac{W^2}{c^2} + m^2 c^2\right)\psi = 0$$

---

* Zusatz bei der Korrektur. In einer inzwischen erschienenen Arbeit (Proc. Roy. Soc. **118**, 351, 1928) hat Herr Dirac den erwähnten divergenzfreien Stromvektor gleichfalls aufgestellt. Da unsere Methode von der seinen verschieden ist, und gleichzeitig eine nähere Analyse der relativistischen Transformations-Eigenschaften der ψ gibt, sind die nachfolgenden diesbezüglichen Ausführungen vielleicht doch nicht ohne Interesse.

** Proc. Roy. Soc. **117**, 610, 1928.

*** Nämlich die von Fock, Gordon, Klein, Kudar und Schrödinger aufgestellte Wellengleichung

$$\sum_{k=1}^{4}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x_k} - \frac{e}{c}\Phi_k\right)^2 + m^2 c^2\Big\}\,\psi = 0,$$

wobei x_1, x_2, x_3 die drei räumlichen Koordinaten sind und $x_4 = ict$ ist, $\Phi_1, \Phi_2, \Phi_3, \Phi_4$ aber das elektromagnetische Viererpotential (Φ_1, Φ_2, Φ_3 reell, Φ_4 rein imaginär). Mit dieser behandelte Gordon den Comptoneffekt, ZS. f. Phys. **40**, 117, 1926; dort wird sie auch näher diskutiert

**** Vgl. Dirac, l. c., bzw. S. 619, 620, 624.

* So z. B., daß der Spin jedem materiellen Gebilde, auch dem Proton, aus relativistischen Gründen zukommen muß.

** Man sieht dies schon daran, daß der Wahrscheinlichkeit (der Koordinatenstatistik) $|\psi|^2$ der Transformationstheorie in der Wellentheorie Schrödingers die elektrische Ladungsdichte entspricht — und diese ist keine relativistische Invariante, sondern die Zeitkomponente eines Vektors.

*** Vgl. Dirac, l. c. S. 610.

$\left(p, q, r, W\right.$ sind die Impuls- und Energieoperatoren $\frac{h}{2\pi i}\frac{\partial}{\partial x}, \frac{h}{2\pi i}\frac{\partial}{\partial y}, \frac{h}{2\pi i}\frac{\partial}{\partial z}, \left.\frac{h}{2\pi i}\frac{\partial}{\partial t}\right)$ ist unbefriedigend, weil sie von zweitem Grade in der Zeit ist, sie kann aber durch die folgende ersetzt werden:

$$\left(\gamma_1 p + \gamma_2 q + \gamma_3 r + \gamma_4 \frac{W}{ci} + m c i\right) \psi = 0,$$

wenn die $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ solche Operatoren sind, die nicht auf x, y, z, t operieren, und den Gleichungen

$$\gamma_k^2 = 1, \quad (k = 1, 2, 3, 4),$$
$$\gamma_k \gamma_l + \gamma_l \gamma_k = 0, \quad (k \neq l, \quad k, l = 1, 2, 3, 4),$$

genügen. Daher muß es weitere Variablen ζ geben, auf die die γ_k operieren. Wenn ζ nur endlich vieler Werte fähig ist, so sind die γ_k Matrizen mit entsprechend viel Zeilen und Kolonnen — die kleinste Zahl, für die die obigen Bedingungen der γ_k erfüllbar sind, ist aber vier. Somit ist ζ vierwertig, und die γ_k vierdimensionale Matrizen.

Dirac zeigte, daß es solche Systeme γ_k gibt, und zwar im wesentlichen nur eines: wenn γ_k ein solches System ist, so entstehen alle anderen derartigen Systeme γ_k' durch

$$\gamma_k' = A^{-1} \gamma_k A, \quad (k = 1, 2, 3, 4),$$

wo A eine beliebige (vierdimensionale) Matrix ist*.

Wenn ein Feld da ist, so sei $\Phi_1, \Phi_2, \Phi_3\ \Phi_4 = i V$ das Viererpotential**, dann gewinnt Dirac seine Gleichung aus der feldfreien durch Ersetzen von p, q, r, W mit

$$p - \frac{e}{c}\Phi_1, \quad q - \frac{e}{c}\Phi_2, \quad r - \frac{e}{c}\Phi_3, \quad W - eV,$$

also auf die übliche Art:

$$\left[\gamma_1\left(p - \frac{e}{c}\Phi_1\right) + \gamma_2\left(q - \frac{e}{c}\Phi_2\right) + \gamma_3\left(r - \frac{e}{c}\Phi_3\right) + \gamma_4 \frac{W - eV}{ci} + mci\right] \Phi = 0^{***}.$$

Von dieser Gleichung wollen wir zunächst die monochromatisch-ebenen Wellenlösungen angeben, d. h. die kräftefrei mit gegebenem

* Vgl. Dirac, l. c. § 3.

** $\Phi_1, \Phi_2, \Phi_3, V$ reell, für $\Phi_1 = \Phi_2 = \Phi_3 = 0$ ist V das gewöhnliche elektrostatische Potential.

*** Wenn man $mci\psi$ auf die linke Seite schafft und quadriert, so kommt man nicht, wie im feldfreien Falle, zur früheren relativistischen Gleichung, es treten vielmehr die Spinglieder auf (l. c. S. 619), die dem „äußeren" magnetischen Felde entsprechen.

Impuls beweglichen Elektronen beschreiben. Dabei wird einiges Licht auf die eingangs erwähnte Schwierigkeit des Ladungsvorzeichens fallen, sowie die dem Spin entsprechende Polarisation der de Brogliewellen (nach C. G. Darwin) in Evidenz gesetzt werden.

III. Die Wellenfunktion ψ hängt von den Variablen x, y, z, t und ζ ab:

$$\psi = \psi\,(x y z t;\ \zeta),$$

wo ζ nur der vier Werte 1, 2, 3, 4 fähig ist; wir werden daher die Abhängigkeit von ihm durch einen Index andeuten:

$$\psi\,(x y z t;\ \zeta) = \psi_\zeta\,(x y z t),$$

derart, daß ψ vier gewöhnliche (Raum-Zeitliche) Wellenfunktionen $\psi_1, \psi_2, \psi_3, \psi_4$ entsprechen. Wir können daher ψ auch als einen von x, y, z, t allein abhängigen Vektor im vierdimensionalen komplexen Raume* auffassen, seine Komponenten sind die $\psi_1, \psi_2, \psi_3, \psi_4$.

Im Falle einer monochromatisch-ebenen Welle mit den Impulsen p_0, q_0, r_0, W_0 ist also

$$\psi = \mathfrak{v} \cdot e^{\frac{2\pi i}{h}(p_0 x + q_0 y + r_0 z + W_0 t)},$$

wo $\mathfrak{v}$ ein (konstanter, komplex-vierdimensionaler) Vektor ist. Wann befriedigt er die kräftefreie Diracsche Gleichung? Es muß offenbar

$$\left(p_0 \gamma_1 + q_0 \gamma_2 + r_0 \gamma_3 + \frac{W_0}{ci} \gamma_4 + mci\,1\right) \mathfrak{v} = 0$$

sein, hier handelt es sich aber um ein rein vierdimensional-geometrisches Problem: $p_0, q_0, r_0, \frac{W_0}{ci}, mci$ sind gewöhnliche Zahlen, $\gamma_1, \gamma_2, \gamma_3, \gamma_4, 1$ Matrizen**.

Wir führen die Matrizen

$$p_0 \gamma_1 + q_0 \gamma_2 + r_0 \gamma_3 + \frac{W_0}{ci} \gamma_4 + mci\,1 = A$$

$$p_0 \gamma_1 + q_0 \gamma_2 + r_0 \gamma_3 - \frac{W_0}{ci} \gamma_4 + mci\,1 = B$$

* Man hüte sich, diesen mit dem reellen, zufällig auch vierdimensionalen $xyzt$-Raume, der „Welt", zu verwechseln!

** Wie $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ gewählt werden (unter Wahrung der Bedingungen aus II), ist zunächst gleichgültig, 1 ist die Einheitsmatrix. Wenn wir von Matrizen und Vektoren schlechthin sprechen, so sind stets komplex-vierdimensionale gemeint; wenn $X = \{x_{kl}\}$ eine Matrix und $\mathfrak{v} = v_1, v_2, v_3, v_4$ ein Vektor ist, so ist $X\mathfrak{v} = \mathfrak{w} = w_1, w_2, w_3, w_4$ der Vektor mit den Komponenten $w_k = \sum_{l=1}^{4} x_{kl} v_l$.

ein, es gilt die $\mathfrak{v}$ mit $A\mathfrak{v} = 0$ zu bestimmen. Hieraus folgt*

$$p_0\gamma_1 + q_0\gamma_2 + r_0\gamma_3 - \frac{W_0}{ci} - mci\,1 = A^\dagger$$

$$p_0\gamma_1 + q_0\gamma_2 + r_0\gamma_3 + \frac{W_0}{ci} - mci\,1 = B^\dagger$$

und somit

$$AB^\dagger = B^\dagger A = A^\dagger B = BA^\dagger = \left(p_0^2 + q_0^2 + r_0^2 - \frac{W_0^2}{c^2} + m^2c^2\right)1.$$

Aus $A\mathfrak{v} = 0$ folgt nun $B^\dagger A\mathfrak{v} = 0$, also $\mathfrak{v} = 0$, wenn

$$p_0^2 + q_0^2 + r_0^2 - \frac{W_0^2}{c^2} + m^2c^2 \neq 0$$

ist — d. h. es gibt dann keine Lösung (außer $\psi \equiv 0$). Wir nehmen daher

$$p_0^2 + q_0^2 + r_0^2 - \frac{W_0^2}{c^2} + m^2c^2 = 0$$

an — d. h. die relativistische Relation zwischen Impuls, Energie und Masse, wie viele linear unabhängige Lösungen $\mathfrak{v}$ (d. h. ψ) gibt es dann?

IV. Wir müssen den Rang μ der Matrix A bestimmen, dann ist die Zahl der linear unabhängigen Lösungen $4 - \mu$.

Sei der Rang von $B^\dagger$ ν. Wegen $AB^\dagger = 0$ ist dann jedenfalls $\mu + \nu \leqq 4$**, wegen $A - B^\dagger = 2mci\,1$ ist $\mu + \nu \geqq 4$***, also $\mu + \nu = 4$. Wenn wir m als Variable ansehen, so hat A fast immer (d. h. mit höchstens endlich vielen Ausnahmen) denselben Rang; aber wenn wir m durch $-m$ ersetzen, so geht A in $B^\dagger$ über, also ist fast immer $\mu = \nu$, d. h. $\mu = \nu = 2$. Für spezielle m können sich μ, ν nur verringern****, aber wegen $\mu + \nu = 4$ ist dies ausgeschlossen. Daher haben $A, B^\dagger$, und somit auch $A^\dagger, B$, stets den Rang 2.

* Wenn $X = \{x_{kl}\}$ ist, so bezeichnen wir mit $X^\dagger$ die konjugiert-transponierte Matrix $X^\dagger = \{\overline{x_{lk}}\}$. Offenbar gilt:

$$(X+Y)^\dagger = X^\dagger + Y^\dagger,\ (XY)^\dagger = Y^\dagger X^\dagger,\ (cX)^\dagger = \bar{c}X^\dagger$$

(c eine Zahl). Wenn $X = X^\dagger$ ist, so ist X hermiteisch, wenn $X^{-1} = X^\dagger$ ist, so ist es unitär. Von $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ nehmen wir an, daß sie hermiteisch sind (sie sind es bei Dirac). Das Symbol † entspricht der „Adjungierten" bei Differential-Operatoren.

** Da $A\mathfrak{v} = 0$ $4 - \mu$ linear unabhängige Lösungen hat, gibt es unter allen $A\mathfrak{v}$ überhaupt genau μ linear unabhängige; da für alle diese $B^\dagger(A\mathfrak{v}) = 0$ ist, ist $\mu \leqq 4 - \nu$, $\mu + \nu \leqq 4$.

*** Unter allen $A\mathfrak{v}$ bzw. $B^\dagger\mathfrak{v}$ gibt es μ bzw. ν linear unabhängige, unter den $\frac{1}{2mci}(A\mathfrak{v} - B^\dagger\mathfrak{v}) = \mathfrak{v}$ also höchstens $\mu + \nu$; dies sind aber alle $\mathfrak{v}$ überhaupt; daher ist $\mu + \nu \geqq 4$.

**** „Zufällige" lineare Abhängigkeiten können auftreten!

Es gibt also zwei lineare unabhängige Lösungen von $A\mathfrak{v} = 0$. Wegen $AB^\dagger = 0$ ist z. B. jede Kolonne von $B^\dagger$ Lösung, und da $B^\dagger$ den Rang 2 hat, gibt es darunter genau zwei linear unabhängige, wir haben also damit alle wesentlichen Lösungen.

Man sieht: die Diracsche Gleichung hat, je nachdem die relativistische Impuls-Energie-Massenrelation erfüllt ist oder nicht, zwei linear unabhängige Lösungen oder keine. Die möglichen (monochromatisch-ebenen) ψ-Wellen sind somit polarisiert, auf eine Weise, die das Auftreten des Spins gerade ermöglicht. Es erhebt sich aber sofort die Frage: Alle Vektoren $\mathfrak{v}$ bilden eine vierdimensionale Gesamtheit, was ist aus den zwei übrigen Dimensionen geworden?

V. Zu gegebenen p_0, q_0, r_0, m gibt es immer zwei W_0, die die Bedingung

$$p_0^2 + q_0^2 + r_0^2 - \frac{W_0^2}{c^2} + m^2c^2 = 0$$

erfüllen: ein $W_0 > 0$ und $-W_0 < 0$. Das erstere ist $\geqq mc^2$, und kommt eigentlich (wie der Grenzübergang $c \to \infty$ zur unrelativistischen Mechanik zeigt) allein in Frage — daß auch das zweite in Erscheinung tritt, ist eben die unangenehme Zweideutigkeit der gegenwärtigen relativistischen Beschreibung, die hier als Unbestimmtheit des Vorzeichens der Energie, d. h. der Richtung des Zeitablaufs, zutage tritt. Wenn wir W_0 durch $-W_0$ ersetzen, geht A in B über, d. h. der zweidimensionale Unterraum (die „Ebene") $A\mathfrak{v} = 0$ im vierdimensionalen Raume in den ebensolchen Unterraum $B\mathfrak{v} = 0$.

Aus $AB^\dagger = 0$ folgt aber, daß diese zwei Ebenen aufeinander senkrecht stehen, d. h. jeder Vektor der einen auf jedem der anderen senkrecht ist*, somit machen sie zusammen den Gesamtraum aus; d. h. jeder Vektor $\mathfrak{v}$ zerfällt eindeutig nach

$$\mathfrak{v} = \mathfrak{v}' + \mathfrak{v}'',\quad A\mathfrak{v}' = 0,\ B\mathfrak{v}'' = 0;$$

* Zwei Vektoren $\mathfrak{v} = v_1, v_2, v_3, v_4$ und $\mathfrak{w} = w_1, w_2, w_3, w_4$ sind senkrecht, wenn ihr „inneres Produkt" verschwindet:

$$(\mathfrak{v}, \mathfrak{w}) = \sum_{k=1}^{4} v_k \overline{w_k} = 0.$$

Es gilt offenbar für alle Matrizen X und alle Vektoren $\mathfrak{v}, \mathfrak{w}$

$$(X\mathfrak{v}, \mathfrak{w}) = (\mathfrak{v}, X^\dagger\mathfrak{w}),\quad (\mathfrak{v}, X\mathfrak{w}) = (X^\dagger\mathfrak{v}, \mathfrak{w}).$$

Die Lösungen von $A\mathfrak{v} = 0$ sind die Spalten von $B^\dagger$ (und deren Linearkombinationen), und die Bedingung $B\mathfrak{w} = 0$ besagt, wie man sich leicht überlegt, daß $\mathfrak{w}$ auf allen Spalten von $B^\dagger$ senkrecht ist — damit ist aber die Behauptung des Textes bewiesen.

$\mathfrak{v}'$ ist der vektorielle Anteil der de Brogliewelle mit dem raumzeitlichen Teile $e^{\frac{2\pi i}{h}(p_0 x + q_0 y + r_0 z + W_0 t)}$, $\mathfrak{v}''$ ebenso mit $e^{\frac{2\pi i}{h}(p_0 x + q_0 y + r_0 z - W_0 t)}$. Die Ebene $A\mathfrak{v} = 0$ entspricht somit positiver Energie und normalen Verhältnissen, die Ebene $B\mathfrak{v} = 0$ negativer Energie und umgekehrtem Zeitablauf.

Lorentz-Invarianz. Transformations-Eigenschaften.

VI. Wenn wir x, y, z, t durch $x_1, x_2, x_3, \frac{1}{ci} x_4$ ersetzen, so können wir die allgemeine Lorentztransformation so schreiben:

$$x'_k = \sum_{l=1}^{4} o_{kl}\, x_l, \quad (k = 1, 2, 3, 4),$$

wo die Matrix $\{o_{kl}\}$ orthogonal ist* und die richtigen Realitätsverhältnisse hat: wenn von k und l keines oder beide $= 4$ sind, so ist o_{kl} reell, wenn genau eins $= 4$ ist, so ist o_{kl} rein imaginär. Daher transformieren sich $x\,y\,z\,t$ in reelle $x'\,y'\,z'\,t'$.

Wenn wir nun ψ' durch

$$\psi'(x'\,y'\,z'\,t') = \psi(x\,y\,z\,t)$$

definieren**, und wenn ψ der Diracschen Gleichung

$$\left[\gamma_1\left(p - \frac{e}{c}\Phi_1\right) + \gamma_2\left(q - \frac{e}{c}\Phi_2\right) + \gamma_3\left(r - \frac{e}{c}\Phi_3\right) + \gamma_4 \frac{W - eV}{ci} + mci\right]\psi = 0$$

genügt (p, q, r, W, wie in II., die Operatoren $\frac{h}{2\pi i}\frac{\partial}{\partial x}, \frac{h}{2\pi i}\frac{\partial}{\partial y}, \frac{h}{2\pi i}\frac{\partial}{\partial t}, \frac{h}{2\pi i}\frac{\partial}{\partial t}$), so genügt ψ', wie man sich leicht überlegt, der Gleichung

$$\left[\gamma'_1\left(p - \frac{e}{c}\Phi'_1\right) + \gamma'_2\left(q - \frac{e}{c}\Phi'_2\right) + \gamma'_3\left(r - \frac{e}{c}\Phi'_3\right) + \gamma'_4 \frac{W - eV}{ci} + mci\right]\psi' = 0.$$

Hier sind $\Phi'_1, \Phi'_2, \Phi'_3, V'$ die relativistisch transformierten Potentiale $\Phi_1, \Phi_2, \Phi_3, V$, und $\gamma'_1, \gamma'_2, \gamma'_3, \gamma'_4$ die folgenden Matrizen:

$$\gamma'_k = \sum_{l=1}^{4} o_{lk}\,\gamma_l, \quad (k = 1, 2, 3, 4).$$

Wegen der Orthogonalität von $\{o_{kl}\}$ ist wieder

$$\gamma'^2_k = 1, \quad (k = 1, 2, 3, 4),$$
$$\gamma'_k\gamma'_l + \gamma'_l\gamma'_k = 0, \quad (k \neq l,\ k, l = 1, 2, 3, 4),$$

und somit existiert eine Matrix A mit

$$\gamma'_k = A^{-1}\gamma_k A.$$

Hieraus folgt aber (wenn wir die Gleichung links mit A multiplizieren):

$$\left[\gamma_1\left(p - \frac{e}{c}\Phi'_1\right) + \gamma_2\left(q - \frac{e}{c}\Phi'_2\right) + \gamma_3\left(r - \frac{e}{c}\Phi_3\right) + \gamma_4 \frac{W - eV'}{ci} + mci\right]A\psi' = 0.$$

Dieses Resultat ist offenbar so zu interpretieren, daß bei derjenigen Änderung des Bezugssystems, die der Lorentztransformation $\{o_{kl}\}$ entspricht, ψ in $A\psi'$ übergeht*. Der Übergang $\psi \to \psi'$ entspricht der Transformation der $xyzt$-Welt, $\psi' \to A\psi'$ hingegen der Transformation des Spins, der Größe ζ.

VII. Die zu $\{o_{kl}\}$ gehörige Lorentztransformation heiße O, dann würde jedem O ein A zugeordnet, und zwar eines, das bis auf einen konstanten Faktor eindeutig festgelegt ist**. Wenn wir verlangen, daß

---

* Komplex orthogonal, d. h.

$$\sum_{l=1}^{4} o_{kl}\, o_{jl} = \sum_{l=1}^{4} o_{lk}\, o_{lj} = \begin{cases}1, & k = j\\ 0, & k \neq j,\end{cases}$$

und nicht unitär, d. h.

$$\sum_{l=1}^{4} o_{kl}\, \overline{o_{jl}} = \sum_{l=1}^{4} o_{lk}\, \overline{o_{lj}} = \begin{cases}1, & k = j\\ 0, & k \neq j.\end{cases}$$

Die Unitarität (und das „innere Produkt" $\sum_{k=1}^{4} v_k \overline{w_k}$) spielt im komplex-vierdimensionalen-Raum der ψ eine Rolle (vgl. S. 873, Anm. *), nicht aber in der $x\,y\,z\,t$-Welt.

** ψ, ψ' als Vektoren gedacht; für die (auch von ζ abhängige) Wellenfunktionen ist

$$\psi'(x'\,y'\,z'\,t';\zeta) = \psi(x\,y\,z\,t;\zeta),$$

d. h. wir transformieren nur die Raum-zeitliche Welt, aber nicht den Spin.

* Dies ist der Diracsche Beweis der Lorentzinvarianz seiner Gleichung (l. c., § 3), den wir der Verständlichkeit des folgenden halber wiederholen mußten.

** Wenn für alle k $A^{-1}\gamma_k A = A'^{-1}\gamma_k A' = \gamma'_k$ ist, so ist AA'^{-1} mit allen γ_k vertauschbar, also wie man leicht einsieht, const. 1.

$\varDelta$ die Determinante 1 hat, so ist dieser Faktor auf die Werte ± 1, $\pm i$ beschränkt. Wenn O und O' die $\varDelta$ und $\varDelta'$ zugeordnet sind, so ist $O \,.\, O'$ offenbar $\varDelta \,.\, \varDelta'$ zugeordnet. Wir haben also in

$$O \rightarrow \varDelta$$

eine (mehrdeutige!) vierdimensionale Darstellung der Lorentzgruppe — bzw. der orthogonalen $\{o_{kl}\}$ mit reellem bzw. reinimaginärem o_{kl} für $k \neq 4$, $l \neq 4$ oder $k = 4$, $l = 4$ bzw. $k = 4$, $l \neq 4$ oder $k \neq 4$, $l = 4$ vor uns.

Die Vermutung läge nahe, daß es die identische ist, d. h.

$$O = \varDelta.$$

Indessen ist dies, wie es aus der weiter unten folgenden Eigenschaft der $\varDelta$ folgt, und man sich auch leicht direkt überzeugt, nicht der Fall: d. h. ψ hat keineswegs die relativistischen Transformationseigenschaften eines gewöhnlichen Vierervektors*.

Wir wollen hier auf die Eigenschaften dieser Darstellung nicht erschöpfend eingehen, nur ein wesentliches Moment sei hervorgehoben: diese vierdimensionale Darstellung zerfällt in zwei zweidimensionale, d. h. bei Einführen eines geeigneten Koordinatensystems im Raume der ψ-Vektoren (oder auch bei geeigneter Wahl der γ_k) transformieren die $\varDelta$ sowohl die zwei ersten als auch die zwei letzten Komponenten von ψ nur untereinander. Oder „koordinatenlos" formuliert: es gibt zwei aufeinander senkrechte zweidimensionale Unterräume unseres vierdimensionalen Vektorraumes, deren jeder durch die $\varDelta$ in sich selbst transformiert wird.

Hierzu genügt es offenbar, eine mit allen $\varDelta$ vertauschbare Matrix mit zwei doppelten Eigenwerten anzugeben: ihre Eigenvektoren** des einen und des anderen Eigenwertes bilden dann die gewünschten Unterräume. Eine solche Matrix ist nun $\gamma_1 \gamma_2 \gamma_3 \gamma_4 = \alpha$.

In der Tat ist

$$\varDelta^{-1} \alpha \varDelta = \varDelta^{-1} \gamma_1 \varDelta \,.\, \varDelta^{-1} \gamma_2 \varDelta \,.\, \varDelta^{-1} \gamma_3 \varDelta \,.\, \varDelta^{-1} \gamma_4 \varDelta$$
$$= \gamma_1' \gamma_2' \gamma_3' \gamma_4' = \prod_{k=1}^{4} \left(\sum_{l=1}^{4} o_{lk} \gamma_l \right),$$

und dies (wegen der Vertauschungseigenschaften der γ_k sowie der Orthogonalität von $\{o_{kl}\}$)

$$= \gamma_1 \gamma_2 \gamma_3 \gamma_4 = \alpha,$$

* Daß eine Größe mit vier Komponenten kein Vierervektor ist, ist ein in der Relativitätstheorie nie vorgekommener Fall, im Diracschen ψ-Vektor tritt das erste Beispiel einer solchen auf. Wieder zeigt sich (vgl. S. 871, Anm. *), wie verschieden der Raum der ψ-Vektoren von der $x\,y\,z\,t$-Welt ist!

** Das heißt Hauptachsenrichtungen.

d. h. α mit $\varDelta$ vertauschbar. Ferner zeigt man leicht

$$\alpha^2 = 1,$$

d. h. α hat die Eigenwerte ± 1, und etwa

$$\gamma_1^{-1} \alpha \gamma_1 = \gamma_1 \alpha \gamma_1 = -\alpha,$$

also hat es dieselben Eigenwerte wie $-\alpha$, d. h. beide doppelt*.

VIII. Die γ_k sind zwar hermiteische Matrizen, die

$$\gamma_k' = \sum_{l=1}^{4} o_{lk} \gamma_l$$

sind es aber, sobald eines der reinimaginären $o_{kl} \neq 0$ ist (d. h. sobald eine Bewegung und nicht bloß Drehung des Bezugssystems erfolgt), nicht mehr, und daher ist dann $\varDelta$ nicht unitär. Somit ändert sich dann beim Übergang $\psi \rightarrow \varDelta \psi'$ die Größe

$$(\psi, \psi) = \sum_{k=1}^{4} |\psi_k|^2,$$

die die „Wahrscheinlichkeit" bzw. die „elektrische Dichte" darstellt. Dies ist (vgl. Anm. **, S. 869) nicht weiter überraschend: diese Größe ist ja relativistisch keine Invariante, sondern die Zeitkomponente eines Vektors. Diesen Vektor gilt es zu bestimmen.

Zu diesem Zwecke bemerken wir folgendes: die Wahrscheinlichkeit (ψ, ψ) ist der Wert der zur Matrix 1 gehörigen Hermiteschen Form, für den Vektor ψ. Anstatt zu fragen, wie sich dies bei Änderungen des Bezugssystems ändert, können wir allgemeiner fragen: wie ändert sich der Wert der zur Matrix U gehörigen Hermiteschen Form für den Vektor ψ?

ψ geht in $\varDelta \psi$ über**, also der Wert unserer Hermiteschen Form $(\psi, U\psi)$ in $(\varDelta \psi, U \varDelta \psi) = (\psi, \varDelta^\dagger U \varDelta \psi)$; wir können also sagen: U geht in $\varDelta^\dagger U \varDelta$ über. Daß die Wahrscheinlichkeit, d. h. 1, nicht invariant ist, besagt eben: es ist allgemein $\varDelta^\dagger \varDelta \neq 1$.

Die Transformationseigenschaft $U \rightarrow \varDelta^\dagger U \varDelta$ ist eine unübersichtliche, angenehmer wäre eine wie $U' \rightarrow \varDelta^{-1} U' \varDelta$, da $\varDelta$ durch etwas Ähnliches definiert wurde***. Für unitäre $\varDelta$ wäre beides dasselbe, aber $\varDelta$ ist nicht unitär!

* α ist Diracs Matrix ϱ_1. Das „geeignete" Koordinatensystem des ψ-Vektorraumes ist offenbar eines, in dem α die Diagonalform hat. Man zeigt übrigens leicht, daß im wesentlichen nur die Matrizen 1 und α mit allen $\varDelta$ vertauschbar sind.

** Von der Abhängigkeit des ψ vom Weltpunkte $x\,y\,z\,t$ bzw. $x'\,y'\,z'\,t'$ sehen wir jetzt ab: ψ sei stets in entsprechenden Weltpunkten genommen.

*** Nämlich

$$\varDelta^{-1} \gamma_k \varDelta = \gamma_k' = \sum_{l=1}^{4} o_{lk} \gamma_l.$$

IX. Trotzdem gilt für A eine ebenso weitgehende (aber andere!) Beschränkung wie die Unitarität, und diese wird uns an unser Ziel bringen; sie werde daher zuerst hergeleitet.

Nach Definition von A ist:

$$A^{-1}\gamma_k A = \gamma'_k = \sum_{l=1}^{4} o_{lk}\gamma_l$$

also nach Anwendung von $\dagger$ (weil die γ_k hermiteisch sind)

$$A^\dagger \gamma_k A^{\dagger -1} = \gamma_k'^\dagger = \sum_{l=1}^{4} \pm o_{lk}\gamma_l,$$

wobei $+$ bzw. $-$ gilt, je nachdem o_{lk} reell oder rein imaginär ist, d. h. ob eine gerade oder ungerade Zahl der k, l gleich 4 ist. D. h. wir müssen im Schema

$$\begin{matrix} o_{11}\gamma_1 + o_{21}\gamma_2 + o_{31}\gamma_3 + o_{41}\gamma_4 \\ o_{21}\gamma_1 + o_{22}\gamma_2 + o_{32}\gamma_3 + o_{42}\gamma_4 \\ o_{31}\gamma_1 + o_{32}\gamma_2 + o_{33}\gamma_3 + o_{44}\gamma_4 \\ o_{41}\gamma_1 + o_{42}\gamma_2 + o_{43}\gamma_3 + o_{44}\gamma_4 \end{matrix}$$

in den drei ersten Zeilen und dann in den drei ersten Spalten das Vorzeichen umkehren. Das letztere können wir bewirken, indem wir rechts und links mit γ_4 multiplizieren, das erstere (da es sich ja um $\gamma'_1, \gamma'_2, \gamma'_3, \gamma'_4$ handelt!), indem wir rechts und links mit γ'_4 multiplizieren*. Also ist

$$A^\dagger\gamma_k A^{\dagger-1} = \gamma_4\gamma'_4\cdot\gamma'_k\cdot\gamma'_4\gamma_4 = \gamma_4\cdot\gamma'_4\gamma'_k\gamma'_4\cdot\gamma_4 = \gamma_4\cdot A^{-1}\gamma_4\gamma_k\gamma_4 A\cdot\gamma_4,$$

d. h.

$$\gamma_4 A\gamma_4 A^\dagger\cdot\gamma_k = \gamma_k\cdot\gamma_4 A\gamma_4 A^\dagger.$$

Da dies für alle γ_k gilt, muß $\gamma_4 A\gamma_4 A^\dagger = 1$ sein**. Diese Gleichung können wir durch linksseitige Multiplikation mit $A^{-1}\gamma_4$ und rechtsseitige mit γ_4 in

$$\gamma_4 A^\dagger\gamma_4 = A^{-1}$$

umformen.

Jetzt sehen wir: wenn U in $A^\dagger U A$ übergeht, so brauchen wir nur $U' = \gamma_4 U$, $U = \gamma_4 U'$ zu setzen; U' geht in

$$\gamma_4\cdot A^\dagger\cdot\gamma_4 U' A = A^{-1} U' A$$

über. U' hat somit die erwünschte Transformationseigenschaft.

* Wegen
$$\gamma_k^2 = \gamma_k'^2 = 1, \qquad (k = 1, 2, 3, 4)$$
$$\gamma_k\gamma_l + \gamma_l\gamma_k = \gamma'_k\gamma'_l + \gamma'_l\gamma'_k = 0, \qquad (k \neq l,\ k,\ l = 1, 2, 3, 4).$$

** Denn $\gamma_4 A\gamma_4 A^\dagger$ ist mit allen γ_k vertauschbar, also gleich const. 1. Da es die Determinante 1 hat, ist die Konstante gleich ± 1, $\pm i$; und da sie von A stetig abhängt und für $A = 1$ gleich 1 ist, ist sie stets gleich 1.

X. Jetzt ist es ein leichtes, mit 1 fertig zu werden: es entspricht bei Zugrundelegung der Transformationseigenschaft $U' \rightarrow A^{-1}U'A$ der Matrix γ_4. Diese geht in $A^{-1}\gamma_4 A = \gamma'_4 = \sum_{l=1}^{4} o_{lk}\gamma_l$ über, ist also die Zeitkomponente des Vektors $\gamma_1, \gamma_2, \gamma_3, \gamma_4$.

Die $\gamma_1, \gamma_2, \gamma_3$ entsprechen analog den Matrizen $\gamma_4\gamma_1, \gamma_4\gamma_2, \gamma_4\gamma_3$, oder wenn wir von den komplex-orthogonalen Vektorkomponenten zu denjenigen übergehen, die die Komponenten des Lorentz-Vektors sind, so wird daraus $ci\cdot\gamma_4\gamma_1$, $ci\cdot\gamma_4\gamma_2$, $ci\cdot\gamma_4\gamma_3$. Die Zahlenwerte gewinnen wir, indem wir in die Hermiteschen Formen dieser Matrizen den Vektor ψ einsetzen.

$$(\psi, ci\cdot\gamma_4\gamma_1\psi),\quad (\psi, ci\cdot\gamma_4\gamma_2\psi),\quad (\psi, ci\cdot\gamma_4\gamma_3\psi)^*.$$

Hierzu kommt die Zeitkomponente (ψ, ψ), diese vier bilden den Strom, welcher (im Gegensatz zu Gordons Strom, vgl. S. 868, Anm.***) die Ableitungen der ψ_k nicht enthält.

Um den Strom mit Hilfe der ψ_k effektiv hinschreiben zu können, müssen wir über die γ_k konkret verfügen. So wird z. B. bei Diracs Wahl der γ_k**

$$(\psi, \psi) = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2$$
$$(\psi, ci\cdot\gamma_4\gamma_1\psi) = 2c\,\Re(\psi_1\overline{\psi_4} + \psi_2\overline{\psi_3})$$
$$(\psi, ci\cdot\gamma_4\gamma_2\psi) = 2c\,\Im(\psi_1\overline{\psi_4} + \psi_2\overline{\psi_3})$$
$$(\psi, ci\cdot\gamma_4\gamma_3\psi) = 2c\,\Re(\psi_1\overline{\psi_3} - \psi_2\overline{\psi_4})$$

(wenn $z = x + iy$ eine komplexe Zahl ist, so ist $\Re z = x$, $\Im z = y$ sein Real- bzw. Imaginärteil).

* Die Matrizen $ci\cdot\gamma_4\gamma_k$ $(k \neq 4)$ sind hermiteisch:

$$(ci\cdot\gamma_4\gamma_k)^\dagger = -ci\cdot\gamma_k\gamma_4 = ci\cdot\gamma_4\gamma_k.$$

** Dort sind $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ bzw. gleich

$$\boxed{\begin{matrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{matrix}}\quad \boxed{\begin{matrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{matrix}}\quad \boxed{\begin{matrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{matrix}}\quad \boxed{\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{matrix}}$$

und daher $i\cdot\gamma_4\gamma_1$, $i\cdot\gamma_4\gamma_2$, $i\cdot\gamma_4\gamma_3$ bzw.

$$\boxed{\begin{matrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{matrix}}\quad \boxed{\begin{matrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{matrix}}\quad \boxed{\begin{matrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{matrix}}$$

XI. Wir müssen vom soeben gefundenen Strome zeigen, daß er stets (für alle reellen Φ_1, Φ_2, Φ_3, V) divergenzfrei ist — sonst ist er kein Äquivalent für den Strom der Gordonschen Theorie. Indessen ist dies leicht direkt zu verifizieren (unter Berücksichtigung der Diracschen Gleichung für ψ).

In der Tat ist:

$$\frac{\partial}{\partial x}(\psi, ci\cdot\gamma_4\gamma_1\psi) + \frac{\partial}{\partial y}(\psi, ci\cdot\gamma_4\gamma_2\psi) + \frac{\partial}{\partial z}(\psi, ci\cdot\gamma_4\gamma_3\psi) + \frac{\partial}{\partial t}(\psi, \psi)^*$$

$$= 2\Re\left\{\left(\psi, ci\cdot\gamma_4\gamma_1\cdot\frac{\partial}{\partial x}\psi\right) + \left(\psi, ci\cdot\gamma_4\gamma_2\cdot\frac{\partial}{\partial y}\psi\right) + \left(\psi, ci\cdot\gamma_4\gamma_3\cdot\frac{\partial}{\partial z}\psi\right) + \left(\psi, \frac{\partial}{\partial t}\psi\right)\right\}$$

$$= \frac{4\pi}{h}\Im\left(\psi, \{ci\cdot\gamma_4\gamma_1\cdot p + ci\cdot\gamma_4\gamma_2\cdot q + ci\cdot\gamma_4\gamma_3\cdot r + W\}\psi\right)$$

$$= -\frac{4\pi c}{h}\Re\left(\psi, \left\{\gamma_4\gamma_1\cdot p + \gamma_4\gamma_2\cdot q + \gamma_4\gamma_3\cdot r + \frac{W}{ci}\right\}\psi\right)$$

$$= -\frac{4\pi c}{h}\Re\left(\psi, \gamma_4\left\{\gamma_1 p + \gamma_2 q + \gamma_3 r + \gamma_4\frac{W}{ci}\right\}\psi\right)$$

$$= -\frac{4\pi c}{h}\Re\left(\psi, \gamma_4\left\{\gamma_1\Phi_1 + \gamma_2\Phi_2 + \gamma_3\Phi_3 + \gamma_4\frac{V}{ci} - 1\,mci\right\}\psi\right)$$

(da Φ_1, Φ_2, Φ_3, V reelle Koordinatenfunktionen sind)

$$= -\frac{4\pi c}{h}\Re\left\{\Phi_1\cdot(\psi, \gamma_4\gamma_1\psi) + \Phi_2\cdot(\psi, \gamma_4\gamma_2\psi) + \Phi_3\cdot(\psi, \gamma_4\gamma_3\psi) - \frac{V}{ci}(\psi, \psi) + mci(\psi, \gamma_4\psi)\right\}$$

$$= -\frac{4\pi c}{h}\Im\left\{\Phi_1\cdot(\psi, i\gamma_4\gamma_1\psi) + \Phi_2\cdot(\psi, i\gamma_4\gamma_2\psi) + \Phi_3\cdot(\psi, i\gamma_4\gamma_3\psi) + \frac{V}{c}(\psi, \psi) + mc(\psi, \gamma_4\psi)\right\}.$$

Und dies verschwindet, denn der Inhalt der Klammer $\{\,\}$ ist reell, da alle Matrizen $i\gamma_4\gamma_1$, $i\gamma_4\gamma_2$, $i\gamma_4\gamma_3$, 1, γ_4 hermiteisch sind.

---

* Es gilt für alle hermiteischen Matrizen X

$$\frac{\partial}{\partial u}(\psi, X\psi) = \left(\frac{\partial}{\partial u}\psi, X\psi\right) + \left(\psi, \frac{\partial}{\partial u}X\psi\right) = \left(X\frac{\partial}{\partial u}\psi, \psi\right) + \left(\psi, X\frac{\partial}{\partial u}\psi\right) = 2\Re\left(\psi, X\frac{\partial}{\partial u}\psi\right).$$

Schlußbemerkungen.

XII. Die soeben angestellten Überlegungen ließen sich weiter verfolgen, insbesondere kann man allgemein feststellen, was für relativistische Transformationseigenschaften beliebigen Matrizen U zukommen können. Da sich U nach $U \to \Lambda^\dagger U \Lambda$ transformiert, ist es zweckmäßig, wieder $U' = \gamma_4 U$ mit der Transformationseigenschaft $U' \to \Lambda^{-1} U' \Lambda$ zu betrachten. Nun sind die 16 Matrizen

$$1;\ \gamma_1, \gamma_2, \gamma_3, \gamma_4;\ \gamma_1\gamma_2, \gamma_1\gamma_3, \gamma_1\gamma_4, \gamma_2\gamma_3, \gamma_2\gamma_4, \gamma_3\gamma_4;\ \gamma_1\gamma_2\gamma_3, \gamma_1\gamma_2\gamma_4, \gamma_1\gamma_3\gamma_4, \gamma_2\gamma_3\gamma_4;\ \gamma_1\gamma_2\gamma_3\gamma_4,$$

wie man leicht zeigt, linear unabhängig, daher ist jedes U' ein lineares Aggregat von ihnen. Andererseits sieht man sofort, daß die fünf Gruppen — in die wir sie durch Semikolon einteilten — die Transformationseigenschaften

Invariante, Vierervektor, Sechservektor (d. i. schiefsymmetrischer Tensor), Vierervektor, Invariante

haben. Nur diese kommen also in Frage.

Der erste Vierervektor entspricht, wie wir sahen, dem Strome, der Sechservektor wohl, was hier nicht näher ausgeführt werde, dem Spin. Was die drei übrigen Größen sind, wissen wir vorläufig nicht.
