QUANTUM THEORY AND FIVE-DIMENSIONAL RELATIVITY THEORY*

Oskar Klein

in Copenhagen, Denmark
Received April 28, 1926

In the following pages I want to point out a simple connection between the proposed theory of Kaluza[1] regarding the connection between electromagnetism and gravitation on one hand and the suggested method of de Broglie[2] and Schrödinger[3] for the treatment of quantum problems on the other hand. Kaluza's theory attempts to connect the ten gravitational potentials g_{ik} of Einstein and the four electromagnetic potentials φ_i with the coefficients γ_{ik} of a line element of a Riemannian space, which besides the four usual dimensions also contains a fifth dimension. The equations of motion of charged particles then take the form of equations of geodesic lines also in electromagnetic fields. When these are interpreted as wave equations by considering the matter as a kind of propagating wave, then one is led almost automatically to a partial differential equation of second order which can be regarded as a generalization of the usual wave equation. If, now, such solutions to these equations are considered in which the fifth dimension appears purely harmonically with a definite period related to Planck's constant, one comes directly to the above-mentioned quantum theoretical methods.

1. Five-Dimensional Theory of Relativity

I begin by giving a short description of the five-dimensional relativity theory which connects closely to Kaluza's theory but differs in some points from it. Consider a five-dimensional Riemannian line element, for which we

*Original in *Z. Phys.* **37** (1926) 895, reproduced here with permission from Springer–Verlag. Translated from the German by Mrs Uta Schuch and Dr. Lars Bergström.

[1] Th. Kaluza, *Sitzungdber. Berl. Akad.* (1921), p. 966.

[2] L. de Broglie, *Ann. Phys.* (10) **3** (1925) 22. Thesis, Paris 1924.

[3] E. Schrödinger, *Ann. Phys.* **79** (1926) 361 and 489.

postulate a meaning independent of the system of coordinates. We write:

$$d\sigma = \sqrt{\sum \gamma_{ik} dx^i dx^k}, \tag{1}$$

where the symbol $\sum$, as everywhere in the following, describes a summation over the doubly appearing indices from 0 to 4. Here $x^0 \ldots x^4$ denote the five coordinates of the space. The 15 quantities γ_{ik} are the covariant components of a five-dimensional symmetric tensor. In order to transform these to the quantities g_{ik} and φ_i of the usual relativity theory we have to make certain special assumptions. First, four of the coordinates, let us say x^1, x^2, x^3, x^4, always have to characterize the usual space–time. Secondly, the quantities γ_{ik} must not depend on the fifth coordinate x^0. From this follows that the allowed coordinate transformations are restricted to the following group[4]:

$$\left.\begin{aligned} x^0 &= x^{0'} + \psi_0(x^{1'}, x^{2'}, x^{3'}, x^{4'}), \\ x^i &= \psi_i(x^{1'}, x^{2'}, x^{3'}, x^{4'}) \quad (i = 1, 2, 3, 4). \end{aligned}\right\} \tag{2}$$

In fact, we should have written a constant times $x^{0'}$ instead of $x^{0'}$. The restriction of the constant to the value unity is, however, quite inessential.

As one can easily show, γ_{00} is invariant under the transformations (2). The assumption $\gamma_{00} =$ const. is therefore allowed. It is tempting to suggest that only the ratios of γ_{ik} have physical significance. In this case this assumption is only a convention that is always possible. Leaving the unit of measure of x^0 indefinite for the time being, we set:

$$\gamma_{00} = \alpha. \tag{3}$$

One shows furthermore that the following differential quantities remain invariant under the transformations (2), namely[4]:

$$d\vartheta = dx^0 + \frac{\gamma_{0i}}{\gamma_{00}} dx^i, \tag{4}$$

$$ds^2 = \left(\gamma_{ik} - \frac{\gamma_{0i}\gamma_{0k}}{\gamma_{00}}\right) dx^i dx^k. \tag{5}$$

[4]Cf. H. A. Kramers, Proc. Amsterdam 23, Nr. 7, 1922, where a discussion of a simple proof for the invariance of $d\vartheta$ and ds^2 is given that is similar to the following considerations.

In these expressions the doubly appearing indices should be summed over from 1 to 4. For such sums we will, as usual, omit the summation sign. The quantities $d\vartheta$ and ds are connected to the line element in the following way:

$$d\sigma^2 = \alpha d\vartheta^2 + ds^2. \tag{6}$$

Because of the invariance of $d\vartheta$ and γ_{00}, it now follows that the four $\gamma_{0i}(i \neq 0)$, if x^0 is held fixed, transform as the covariant components of an ordinary four-vector. If x^0 is also transformed, there appears additively the gradient of a scalar. This means that the quantities:

$$\frac{\partial \gamma_{0i}}{\partial x^k} - \frac{\partial \gamma_{0k}}{\partial x^i},$$

transform as the covariant components F_{ik} of the electromagnetic field tensor. The quantities γ_{0i} are thus from the point of view of invariance theory behaving as the electromagnetic potentials φ_i. Therefore we assume

$$d\vartheta = dx^0 + \beta\varphi_i dx^i, \tag{7}$$

that is

$$\gamma_{0i} = \alpha\beta\varphi_i \quad (i = 1, 2, 3, 4), \tag{8}$$

where β is a constant and where the φ_i are so defined that in orthogonal Galilean coordinates:

$$\left.\begin{aligned} (\varphi_x, \varphi_y, \varphi_z) &= A, \\ \varphi_t &= -cV, \end{aligned}\right\} \tag{9}$$

where A is the ordinary vector potential, V the ordinary scalar potential and c represents the speed of light.

We want to identify the differential ds with the line element of the usual standard relativity theory. We thus set

$$\gamma_{ik} = g_{ik} + \alpha\beta^2\varphi_i\varphi_k, \tag{10}$$

where we want to choose g_{ik} so that in orthogonal Galilean coordinates:

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2. \tag{11}$$

Hereby the quantities γ_{ik} are brought back to known quantities. The problem is now to find such field equations for the quantities γ_{ik} that g_{ik} and φ_i for sufficient accuracy are given by the field equations of the standard relativity theory. We do not want to examine this difficult problem further here but we want to show that the ordinary field equations can be easily embraced from the viewpoint of the five-dimensional geometry. We form the invariant:

$$P = \sum \gamma^{ik} \left[\frac{\partial \begin{Bmatrix} i\mu \\ \mu \end{Bmatrix}}{\partial x^k} - \frac{\partial \begin{Bmatrix} ik \\ \mu \end{Bmatrix}}{\partial x^\mu} + \begin{Bmatrix} i\mu \\ \nu \end{Bmatrix} \begin{Bmatrix} k\nu \\ \mu \end{Bmatrix} - \begin{Bmatrix} ik \\ \mu \end{Bmatrix} \begin{Bmatrix} \mu\nu \\ \nu \end{Bmatrix} \right], \tag{12}$$

where γ^{ik} are the contravariant components of a five-dimensional fundamental metric tensor and where $\begin{Bmatrix} rs \\ i \end{Bmatrix}$ represents the Christoffel three-index symbol, that is

$$\begin{Bmatrix} rs \\ i \end{Bmatrix} = \frac{1}{2} \sum \gamma^{i\mu} \left(\frac{\partial \gamma_{\mu r}}{\partial x^s} + \frac{\partial \gamma_{\mu s}}{\partial x^r} - \frac{\partial \gamma_{rs}}{\partial x^\mu} \right). \tag{13}$$

In the expression for P we have in mind that all the quantities are independent of x^0 and that $\gamma_{00} = \alpha$.

Let us now consider the integral, evaluated over a closed area of the five-dimensional space:

$$J = \int P \sqrt{-\gamma} dx^0 dx^1 dx^2 dx^3 dx^4, \tag{14}$$

where γ stands for the determinant of the γ_{ik}. We form δJ by varying the quantities γ_{ik} and $\frac{\partial \gamma_{ik}}{\partial x^l}$ where their boundary values are not to be changed. Here α should be considered to be a constant. The variational principle

$$\partial J = 0, \tag{15}$$

then leads to the following equations:

$$R^{ik} - \frac{1}{2} g^{ik} R + \frac{\alpha \beta^2}{2} S^{ik} = 0 \quad (i, k = 1, 2, 3, 4), \tag{16a}$$

and

$$\frac{\partial \sqrt{-g} F^{i\mu}}{\partial x^\mu} = 0 \quad (i = 1, 2, 3, 4), \tag{16b}$$

where R represents Einstein's curvature invariant, R^{ik} the contravariant components of Einstein's curvature tensor, g^{ik} the contravariant components of Einstein's fundamental tensor, S^{ik} the contravariant components of the electromagnetic energy-momentum tensor, g the determinant of the g_{ik} and finally $F^{i\mu}$ the contravariant components of the electromagnetic field tensor. If we set

$$\frac{\alpha\beta^2}{2} = \kappa, \tag{17}$$

where κ represents the gravitational constant used by Einstein, we see that the equations (16a) are in fact identical with the equations of relativity theory for the gravitational field, and (16b) are identical with the generalized Maxwell's equations of relativity theory for a matter-free field point.[5]

If we restrict ourselves to the usual schematic way of treating matter in electron theory and relativity theory, we can obtain the usual equations for the non-matter-free case in a similar way. We replace P in (14) by

$$P + \kappa \sum \gamma_{ik}\Theta^{ik}.$$

In order to define the Θ^{ik}, we first want to consider the tensor appropriate for an electron or a hydrogen nucleus:

$$\vartheta^{ik} = \frac{dx^i}{dl}\frac{dx^k}{dl}, \tag{18}$$

where the dx^i represent the changes of position of the particle, and dl is a certain invariant differential. The Θ^{ik} should be equal to the sums of the ϑ^{ik} for the different particles, per unit volume. We then get back to equations of the ordinary type which become identical to the ordinary field equations, if we set:

$$v_0\frac{d\tau}{dl} = \pm\frac{e}{\beta c}, \tag{19}$$

$$\frac{d\tau}{dl} = \begin{cases}\sqrt{M} \\ \sqrt{m}\end{cases}, \tag{20}$$

[5] See, e.g., B. W. Pauli, *Relativitätstheorie*, pp. 719 and 724.

where in general

$$v_i = \sum \gamma_{i\mu} \frac{dx^\mu}{dl} \tag{21}$$

are the covariant components of the five-dimensional velocity vector v^i, where

$$v^i = \frac{dx^i}{dl}. \tag{22}$$

Further, e represents the electric elementary quantum, M and m the masses of the hydrogen nucleus and electron, respectively. Here the upper symbol pertains to the nucleus, the lower to the electron. In addition,

$$d\tau = \frac{1}{c}\sqrt{-ds^2}$$

is the differential of proper time.

From the field equations, there follow naturally in an ordinary way the equations of motion for the matter particles and the continuity equation. The calculations which lead to this can be easily summarized from our point of view. As one can easily see, our field equations are, namely, equivalent to the following 14 equations:

$$P^{ik} - \frac{1}{2}\gamma^{ik}P + \kappa\Theta^{ik} = 0 \tag{23}$$

($i, k = 0, 1, 2, 3, 4,$ but not both of them zero), where the P^{ik} are the contravariant components of the reduced five-dimensional curvature tensor (corresponding to the R^{ik}). The equations in question now follow by forming the divergence of (23). From this follows that the charged particles move on five-dimensional geodesic lines which satisfy the conditions (19) and (20).[6] As one sees immediately, these conditions are therefore compatible with the equations of geodesic lines because x^0 does not appear in γ_{ik}.

Here must be recalled that there really do not exist sufficient reasons for the exact validity of Einstein's field equations. Nevertheless, it might

[6]The special values of $\frac{d\tau}{dl}$ are of course in this connection without significance. Important is here only $\frac{d\tau}{dl} = \text{const.}$

not be without interest that all the 14 field equations can be summarized in such an easy way from the point of view of Kaluza's theory.

2. The Wave Equation of the Quantum Theory

We are now going to relate the theory of stationary states, and the corresponding characteristic deviations from the mechanics which appear in the modern quantum theory, to the five-dimensional theory of relativity. Let us consider for this purpose the following differential equation which should be related to our five-dimensional space and which can be considered as a simple generalization of the wave equation:

$$\sum a^{ik}\left(\frac{\partial^2 U}{\partial x^i \partial x^k} - \sum \left\{ {ik \atop r} \right\} \frac{\partial U}{\partial x^r}\right) = 0. \tag{24}$$

Here the a^{ik} signify the contravariant components of a five-dimensional symmetric tensor which should be certain functions of the coordinates. The equation (24) is valid independently of the coordinate system.

Let us first consider a wave propagation determined by (24) which corresponds to the limiting case of geometrical optics. We arrive at this if we set:

$$u = Ae^{i\omega\Phi}, \tag{25}$$

and assume ω so large that in (24) only those terms have to be taken into account that are proportional to ω^2. We then obtain:

$$\sum a^{ik} \frac{\partial \Phi}{\partial x^i} \frac{\partial \Phi}{\partial x^k} = 0, \tag{26}$$

an equation which corresponds to the Hamilton–Jacobi partial differential equation of mechanics. If we set:

$$p_i = \frac{\partial \Phi}{\partial x^i}, \tag{27}$$

the differential equations for the rays, as is well known, can be written in the following Hamiltonian form:

$$\frac{dp_i}{-\dfrac{\partial H}{\partial x^i}} = \frac{dx^i}{\dfrac{\partial H}{\partial p_i}} = d\lambda, \tag{28}$$

where

$$H = \frac{1}{2} \sum a^{ik} p_i p_k. \tag{29}$$

From (26) there also follows that

$$H = 0. \tag{30}$$

A different representation of this equation which corresponds to the Lagrangian form, follows through the circumstance that the rays can be regarded as geodesic zero lines of the differential form:

$$\sum a_{ik} dx^i dx^k,$$

where the a_{ik} represent reciprocal quantities to the a^{ik}, that is

$$\sum a_{i\mu} a^{k\mu} = \delta_i^k = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}. \tag{31}$$

If we now set

$$\sum a_{ik} dx^i dx^k = \mu (d\vartheta)^2 + ds^2, \tag{32}$$

we can achieve, by an appropriate choice of the constant μ, that our ray equations become identical to the equations of motion of charged particles. If we set, in order to see this:

$$L = \frac{1}{2}\mu \left(\frac{d\vartheta}{d\lambda}\right)^2 + \frac{1}{2}\left(\frac{ds}{d\lambda}\right)^2, \tag{33}$$

there follows

$$p_0 = \frac{\partial L}{\partial \dfrac{dx^0}{d\lambda}} = \mu \frac{d\vartheta}{d\lambda}, \tag{34}$$

and

$$p_i = \frac{\partial L}{\partial \dfrac{dx_i}{d\lambda}} = u_i \frac{d\tau}{d\lambda} + \beta p_0 \varphi_i \quad (i = 1, 2, 3, 4), \tag{35}$$

where $u_1 \ldots u_4$ represent the covariant components of the ordinary velocity vector. The ray equations now have the form:

$$\frac{dp_0}{d\lambda} = 0, \tag{36a}$$

$$\frac{dp_i}{d\lambda} = \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^i}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} + \beta p_0 \frac{\partial \varphi_\mu}{\partial x^i}\frac{dx^\mu}{d\lambda} \quad (i = 1, 2, 3, 4). \tag{36}$$

From

$$\mu d\vartheta^2 + ds^2 = \mu d\vartheta^2 - c^2 d\tau^2 = 0,$$

there follows

$$\mu \frac{d\vartheta}{d\tau} = c\sqrt{\mu}. \tag{37}$$

Since, according to (34) and (36a), $\frac{d\vartheta}{d\lambda}$ and therefore also $\frac{d\tau}{d\lambda}$ are constant, we can choose λ such that

$$\frac{d\tau}{d\lambda} = \begin{cases} M \text{ for the hydrogen nucleus} \\ m \text{ for the electron.} \end{cases} \tag{38}$$

Furthermore, in order to get to the ordinary equations of motion, we have to assume:

$$\beta p_0 = \begin{cases} +\frac{e}{c} \text{ for the hydrogen nucleus} \\ -\frac{e}{c} \text{ for the electron.} \end{cases} \tag{39}$$

From (37) there follows then:

$$\mu = \begin{cases} \frac{e^2}{\beta^2 M^2 c^4} \text{ for the hydrogen nucleus} \\ \frac{e^2}{\beta^2 m^2 c^4} \text{ for the electron.} \end{cases} \tag{40}$$

The equations (35) and (36) then completely agree with the ordinary equations of motion of charged particles in gravitational fields and electromagnetic fields. In particular, the quantities p_i, defined according to (35), are identical with the generalized momenta defined in the usual way, which is important for the following considerations. Since we can still choose β arbitrarily, we will set:

$$\beta = \frac{e}{c}. \tag{41}$$

It then follows simply that

$$p_0 = \begin{cases} +1 \text{ for the hydrogen nucleus} \\ -1 \text{ for the electron.} \end{cases} \tag{39a}$$

$$\mu = \begin{cases} \dfrac{1}{M^2c^2} \text{ for the hydrogen nucleus} \\ \dfrac{1}{m^2c^2} \text{ for the electron.} \end{cases} \tag{40a}$$

As one sees, for the square root in (37) we have to choose the positive sign in the case of the nucleus and the negative sign in the case of the electron. This is indeed rather unsatisfactory. But the fact that one obtains for a single value of μ two different classes of rays which in a certain way are related to each other like positively and negatively charged particles, could be understood as a hint that it might be possible to change the wave equation so that the equations of motion for both kinds of particles follow from a single set of values of the coefficients. We do not now want to enter into this question further, but are going to consider more closely the wave equation that follows from (32) in the case of the electron.

Since for the electron it was assumed that $p_0 = -1$, as a consequence of (27) we have to set:

$$\Phi = -x^0 + S(x^1, x^2, x^3, x^4). \tag{42}$$

De Broglie's theory now follows if we look for the standing waves compatible with the wave equation corresponding to a certain value of ω and thereby assume that the wave propagation proceeds according to the laws of geometrical optics. For that purpose we need the well-known law of the conservation of phase, which immediately follows from (28) and (30). Namely, it follows that

$$\frac{d\Phi}{d\lambda} = \sum \frac{\partial \Phi}{\partial x^i}\frac{dx^i}{d\lambda} = \sum p_i \frac{\partial H}{\partial p_i} = 2H = 0. \tag{43}$$

The phase is thus carried along by the wave. Let us now consider the simple case where Φ can be split into two parts, one of which depends only on a single coordinate, let us say x, which swings back and forth periodically with time. Then, a standing wave will be possible, which is characterized by the fact that a harmonic wave represented at a certain moment by (25) after one period in x meets in phase with that wave which results from the same solution (25) by inserting the new values of x^0, x^2, x^3, x^4. Because of the conservation of phase, the condition for this is simply:

$$\omega \oint p dx = n 2\pi, \tag{44}$$

where n is an integer. If we set:

$$\omega = \frac{2\pi}{h}, \tag{45}$$

where h represents Planck's constant, the ordinary quantization condition for a separable coordinate then results. The analogous situation is of course true for an arbitrary periodic system. The ordinary quantum theory of periodic systems thus corresponds completely to the treatment of interference phenomena through the assumption that the waves propagate according to the laws of geometrical optics. It may also be emphasized that because of (42), the relations (44), (45) are invariant under the coordinate transformations (2).

Let us now also consider the equation (24) in the case where ω is not so large so that we only have to take into account terms to the second power in ω. We thereby restrict ourselves to the simple case of an electrostatic field. We then have in Cartesian coordinates:

$$\left.\begin{aligned} d\vartheta &= dx^0 - eV\, dt, \\ ds^2 &= dx^2 + dy^2 + dz^2 - c^2 dt^2. \end{aligned}\right\} \tag{46}$$

Therefore there follows

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{1}{2c^2}(p_t + eVp_0)^2 + \frac{m^2c^2}{2} p_0^2. \tag{47}$$

In equation (24) we can neglect the quantities that are proportional to $\left\{ {ik \atop \tau} \right\}$, since according to (17) the three-index symbols are in this case small quantities proportional to the gravitational constant κ. We therefore obtain[7]

$$\Delta U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} - \frac{2eV}{c^2}\frac{\partial^2 U}{\partial t \partial x^0} + \left(m^2c^2 - \frac{e^2V^2}{c^2}\right)\frac{\partial^2 U}{\partial x^{02}} = 0. \tag{48}$$

Since V depends only on x, y, z, we can set for U in agreement with (42) and (45):

$$U = e^{-2\pi i\left(\frac{x^0}{h} - \nu t\right)}\psi(x, y, z). \tag{49}$$

This inserted in (48), yields

$$\Delta\psi + \frac{4\pi^2}{c^2h^2}[(h\nu - eV)^2 - m^2c^4]\psi = 0. \tag{50}$$

If we set additionally:

$$h\nu = mc^2 + E, \tag{51}$$

we obtain the equation given by Schrödinger,[8] whose standing waves correspond, as known, to the values of E which are identical to the energy values calculated from Heisenberg's quantum theory. As one sees, E is in the limiting case of geometrical optics equal to the mechanical energy defined in the usual way. As Schrödinger emphasized, the frequency condition says, according to (51), that the light frequencies belonging to the system are equal to the differences that are formed from the different values of the frequency ν.

3. Final Remarks

Like the papers of de Broglie, the above considerations have arisen from the endeavour to use the analogy between mechanics and optics, which

[7]Except for the appearance of x^0, which is negligible for this application, this equation differs from the Schrödinger equation by the way in which the time appears in (48). In support of this form of the quantum equation one can mention, in the case where V depends harmonically on time, that this equation has solutions which correspond to the dispersion theory of Kramers as do Schrödinger's solutions to the quantum theory of spectral lines, which can be shown by a simple perturbation calculation. I owe this remark to Dr. W. Heisenberg.
[8]Schrödinger, Ref. 3.

appears in the Hamiltonian method, for a deeper understanding of quantum phenomena. The similarity of conditions for the stationary states of an atomic system to the interference phenomena of optics indeed seems to indicate that this analogy has a real physical significance. Now, concepts such as point charge and material point are indeed strange in classical field physics. Of course, the hypothesis has often been maintained that the matter particles have to be interpreted as special solutions of field equations, which determine the gravitational field and the electromagnetic field. It is tempting to relate the mentioned analogy to this concept. Because according to this hypothesis, it is really not so strange that the motion of material particles has similarities to the propagation of waves. The analogy under discussion, however, is incomplete, as long as one considers the wave propagation in a space of only four dimensions. This already appears in the variable speed of material particles. But if one imagines, however, the observed motion as a kind of projection on the space–time of a wave propagation, which proceeds in a space of five dimensions we can, as we saw, make the analogy complete. In mathematical terms this means that the Hamilton–Jacobi equation cannot be interpreted as a characteristic equation of a four-dimensional but rather of a five-dimensional wave equation. In this way one is led to Kaluza's theory.

Although the introduction of a fifth dimension in our physical considerations might seem strange at the outset, a radical modification of the geometry based upon the field equations is suggested in a totally different way by the quantum physics. For, as is well known, it is less and less probable that the quantum phenomena allow a unified space–time description, whereas the possibility of describing these phenomena by a system of five-dimensional field equations probably cannot be excluded a priori.[9] Whether there is something real behind these indications of possibilities, the future of course will have to decide. In any case it must be emphasized that the way of treatment attempted in this note, concerning the field equations as well as the theory of stationary states, has to be regarded as provisional. This is in particular true for the schematic way of treating matter, mentioned on page 71, as well as the circumstance that the two kinds

[9] Remarks of this kind, which Prof. Bohr has made on several occasions, have had a decisive influence on the creation of the present note.

of charged particles are treated by different equations of Schrödinger's type. The question is also left quite open whether one can be content with the 14 potentials when describing physical processes, or whether the Schrödinger method means the introduction of a new state quantity.

I have been occupied with the considerations presented in this note at the Physical Institute of the University of Michigan, Ann Arbor, as well as at the here present institute for theoretical physics. At this point I also want to express my warmest thanks to Prof. H. M. Randall and Prof. N. Bohr.