"Zur kovarianten Formulierung der Diracschen Gleichung," Zeit. f. Physik **57** (1929), 474-483.

On the covariant formulation of the Dirac equation

By **Cornel Lanczos** in Berlin

(Received on 3 August 1929)

Translated by D. H. Delphenich

As a continuation of the previous investigation, the formalism that was applied there is extended to the case where an external field is present, and the covariant formulation will again be carried out. The system of equations that is obtained admits no restriction in the manifold of quantities if an undesirable coincide is to be avoided. The mixed appearance of the dual object proves to be an intrinsic unreality of the system and leaves one to conjecture a modification of the Dirac equation that lies at the basis for it.

In a previous paper [*] we showed that the Dirac system for the free electron has a close connection with Hamilton's quaternion operator, and as a result of the close relationship of that operator to the tensor-analytic picture of four-dimensional space, we arrived at a complete tensor-analytic description. The operator, when given an imaginary unit, thus proved to be a purely formal tool, but it vanishes for the resulting system that still contains only real vector-analytic quantities. In particular, an anti-symmetric tensor of the same type as an electromagnetic field strength emerged, along with two vectors that are to be set parallel to the electric (magnetic, resp.) current of the Maxwell theory. In fact, the coupling of these quantities corresponded to the Maxwell equations completely, except that yet another coupling is present that, in conjunction with other ones, implies the appearance of the Schrödinger equation. Since we have been concerned with only the equation for the free electron up to now, the examination shall now be extended to the case in which an external electric field is also under consideration.

The prescription for the extension of the Dirac equation reads as follows: One extends the operation $\partial / \partial x_i$ by the addition of $i\Phi_i$, where Φ_i means the external vector potential multiplied by $2\pi e / h$:

$$\Phi_i = \frac{2\pi}{h} e\varphi_i . \tag{1}$$

It is obvious that this extension is also possible in our covariant system by the same prescription without disturbing the covariant. The components Φ_i then transform just like the $\partial / \partial x_i$, and the appearance of i introduces no complications, since arbitrary complex quantities may appear in our quaternion schema; nevertheless, the real interpretation is inherently insured.

[*] ZS. f. Phys. **57**, 447, 1929; in the following, this will be cited as *loc. cit.*

We once more write down our basic equations (54), *loc. cit.*, and indeed in a form that makes no use of the specific reality properties of the actual spacetime continuum, and thus insures the actual covariance:

$$\left.\begin{aligned} \bar{\nabla} F &= \alpha G^* \\ \bar{\nabla} G^* &= \alpha F \end{aligned}\right\} \tag{2}$$

If we now apply the rule that the operation $\partial / \partial x_i$ is to be replaced with $\partial / \partial x_i + i\Phi_i$ then the following system results:

$$\left.\begin{aligned} \bar{\nabla} F + i\bar{\Phi} F &= \alpha G^*, \\ \bar{\nabla} G^* + i\Phi G^* &= \alpha F. \end{aligned}\right\} \tag{3}$$

Φ is therefore the quaternion that represents the vector potential Φ_i whose components are then the Φ_i . Its spatial part is real, while its temporal part is imaginary, and one thus has – precisely as one has for ∇ – the rule:

$$\Phi^* = -\bar{\Phi}. \tag{4}$$

We now once more write equation (3) in such a form that we take the complex conjugate in the second equation, and thus obtain our basic equations with the desired extension in the following form:

$$\left.\begin{aligned} \bar{\nabla} F + i\bar{\Phi} F &= \alpha G^*, \\ \bar{\nabla} G - i\bar{\Phi} G &= -\alpha F^*. \end{aligned}\right\} \tag{5}$$

Before we go on to the consideration of the total system of equations, we would first like to look into the Dirac equation that a subgroup of the total system must represent. We saw [cf., equation (61), *loc. cit.*] that one such exists for the combination:

$$H = \tfrac{1}{2}[(F + G) + i(F - G)\, j_z], \tag{6}$$

just as it does for the combination:

$$H' = \tfrac{1}{2}[(G + F) + i(G - F)\, j_z]. \tag{6'}$$

By adding (subtracting, resp.) of equations (5), we obtain:

$$\left.\begin{aligned} \bar{\nabla}(H + H') - \bar{\Phi}(H - H')\, j_z &= \alpha i (H - H')^*\, j_z, \\ \bar{\nabla}(H - H') - \bar{\Phi}(H + H')\, j_z &= \alpha i (H + H')^*\, j_z, \end{aligned}\right\} \tag{7}$$

and this yields, by adding (subtracting, resp.):

$$\left.\begin{aligned} \bar{\nabla} H - \bar{\Phi} H j_z &= \alpha i H^*\, j_z, \\ \bar{\nabla} H' - \bar{\Phi} H' j_z &= -\alpha i H'^*\, j_z. \end{aligned}\right\} \tag{8}$$

These are two Dirac equations for H (H', resp.) alone that are obviously equivalent to the original system (5). One easily convinces oneself that the introduction of the vector potential, in fact, consists of the replacement of the operation $\partial / \partial x_i$ by $\partial / \partial x_i + i\Phi_i$. One then immediately recognizes from the association (60), *loc. cit.*, that the multiplication of ψ by i is equivalent to a multiplication by $-j_z$ for H. It is remarkable that in the two equations the vector potential emerges with the opposite sign, such that the operation $\partial / \partial x_i + i\Phi_i$ substitutes in the first equation while the operation $\partial x_i - i\Phi_i$ substitutes in the second one.

It is of interest to examine the transformations properties of the completed Dirac system. We again perform a Lorentz transformation [cf., equation (18), *loc. cit.*], in which we set:

$$\left.\begin{aligned} \nabla' &= p \nabla \bar{p}^*, \\ \Phi' &= p \Phi \bar{p}^*, \end{aligned}\right\} \tag{9}$$

and again let the transformation of H be [cf., equation (85) itself]:

$$H' = p\, H\, k. \tag{10}$$

[This H' has nothing to do with the H' of equations (8).]

We now obtain two conditions for the quaternion k, namely:

$$\left.\begin{aligned} k^* j_z &= j_z k, \\ k\, j_z &= j_z k. \end{aligned}\right\} \tag{11}$$

The second condition comes about from the term with the vector potential and implies that k must be real:

$$k = k^*. \tag{12}$$

Then, however, k can only have a j_z component and a j_l component, so there only two degrees of freedom left, which can be reduced to a single one by a length normalization. However, this implies only a phase transformation of the ψ, which is indeed quantum-mechanically allowable.

Therefore, any of the transformation groups that we dealt with in the previous paper (see, section 9) are obsolete, and the usual transformation theory of ψ quantities remains valid. Moreover, the objection that was raised against the current vector of the Dirac theory is then no longer justified, and the covariance of the current vector is, in fact, insured.

It then seems all the more noteworthy that under the covariant extension of the system any quaternion:

$$F\bar{F}^* + G\bar{G}^*, \tag{13}$$

that is composed of F and G, corresponds to the current vector, and is divergence-free does not represent a vector here, and can possess no vector-analytic sense whatsoever. (The vector character is present only for purely spatial rotation.) [†]

For the derivation of the Schrödinger wave equation, we now apply the operator ∇ to the first of equations (5):

$$\begin{aligned}\nabla\bar{\nabla}F + i\nabla(\bar{\Phi}F) &= \alpha\nabla G^* = \alpha(\alpha F - i\Phi\, G^*) \\ &= \alpha^2 F - i\Phi\bar{\nabla}F + \Phi\bar{\Phi}F\,. \end{aligned} \tag{14}$$

However, one has:

$$\nabla(\bar{\Phi}F) = (\nabla\bar{\Phi})F + \overline{\Phi(\bar{\nabla}\cdot F)}\,, \tag{15}$$

and if we add the terms on the right-hand side then we see that – with the exception of the characteristic term $(\nabla\bar{\Phi})F$, which is known to represent the electron spin for the Dirac theory – all other operations have a scalar character. Obviously, one has, in fact:

$$\Phi\bar{\nabla} + \overline{\Phi\bar{\nabla}} = 2\,\Phi_\nu \frac{\partial}{\partial x_\nu}, \tag{16}$$

and we can thus write equation (14) as follows:

$$\left(\frac{\partial^2}{\partial x_\nu^2} + 2i\Phi_\nu \frac{\partial}{\partial x_\nu} - \alpha^2 - \Phi_\nu^2\right)F = -\,i(\nabla\bar{\Phi})F\,. \tag{17}$$

Exactly the same equation is also true for G, except that the i must be replaced by – i (Φ by – Φ, resp.).

Equation (17) represents the Schrödinger wave equation, extended by the spin interaction, which emerges as a specific achievement of the Dirac theory with no special assumptions.

We now leave behind the quaternion formalism, and we would like to write down our equations in the tensor-analytic interpretation, where F is again an anti-symmetric tensor (whose temporal part is invariant), so G should be regarded as a complex vector. We employ the same relationships as in our first treatise [see, equations (96)], except that now we would like to propose a notation that avoids the appearance of imaginary quantities entirely. To this end, we abandon the four Minkowskian coordinates and think of the real time as having been introduced as a fourth coordinate. We then have only to make the distinction between covariant and contravariant, although, at the same time, we can also bring the system into a form in which it remains invariant under not only linear, but arbitrary point transformations, so it is in a generally covariant form. Our system of equations thus relates to arbitrary curvilinear coordinates.

[†] Remark by the editor: The simple and fundamental meaning of this object, namely, that it represents the energy current that, combined with the impulse current extends to a tensor of second order, was first made known to the author in connection with this investigation. Cf., the paper: "Die Erhaltungssätze in der feldmässigen Darstellung der Diracschen Theorie" that follows next in this journal.

On the one hand, we have a system (A), which is analogous to the Maxwell equations, and must now be modified as follows:

$$\left.\begin{aligned}\frac{\partial S}{\partial x_\nu}g^{i\nu}+\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}F^{i\nu}}{\partial x_\nu}&=\alpha S^i-\Phi_\nu\tilde{F}^{i\nu}-\Phi^i M,\\ \frac{\partial M}{\partial x_\nu}g^{i\nu}+\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}\tilde{F}^{i\nu}}{\partial x_\nu}&=\alpha M^i+\Phi_\nu F^{i\nu}+\Phi^i S.\end{aligned}\right\}(A)\qquad(18A)$$

One adds the following "back-reaction system":

$$\left.\begin{aligned}\frac{\partial S_i}{\partial x_k}-\frac{\partial S_k}{\partial x_i}+\left(\frac{\partial M_i}{\partial x_k}-\frac{\partial M_k}{\partial x_i}\right)&=\alpha F_{ik}+(\Phi_i M_k-\Phi_k M_i)-(\Phi_i S_k-\Phi_k S_i),\\ \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}M^\nu}{\partial x_\nu}&=\alpha M+\Phi_\nu S^\nu,\\ \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}S^\nu}{\partial x_\nu}&=\alpha S-\Phi_\nu M^\nu.\end{aligned}\right\}(B)\qquad(18B)$$

Therefore, g_{ik} refers to the metric tensor, and g, to the determinant of the metric (the determinant multiplied by – 1, resp., due to the negative index of inertia of the spacetime line elements.)

The dual association is takes the form of the following schema for an anti-symmetric tensor of second order:

$$\tilde{F}^{12}=\frac{1}{\sqrt{g}}F_{34},\text{ etc.,}\tag{19}$$

$$\tilde{F}_{12}=\sqrt{g}\,F^{34},\text{ etc.}\tag{19′}$$

If we compare the system thus obtained with the one in the previous paper, where the vector potential did not appear (section 10) then this yields a number of remarkable distinctions that seem to stand in the way of a simple interpretation of the equations to a large degree, which we originally suspected. One can regard the previous schema as a generalization and extension of the Maxwell equations. The generalization consists in the appearance of two scalars S and M that enter in, the fact that the system includes no intrinsic dependencies, and the fact that the conservation law that is obtained for electric and magnetic current does not follow as a necessary consequence here. The extension consists in the addition of two new equations, which we have referred to as "back-coupling," because we regard them as the interaction of the system with itself. The manifold of quantities that appears can thus be restricted when we set the two scalars equal to zero, as well as the magnetic current, and this arrive at a closer connection with the usual Maxwell schema. Although we then have 16 equations for 10 functions (viz., the field strengths and the electric current), no coincidence comes about, since the surplus equations would be fulfilled as a consequence of the remaining ones, which are already fulfilled by themselves.

Here, we seek to do something similar, so we see that nothing can be set to zero, without an agreement following as a consequence. If we were to set, say, S and M equal to zero then this would yields that the last two equations of system (B) can be a consequence of the remaining ones only when Φ_i is constant, since otherwise an addition term would appear that depends upon the external field strengths and does not vanish. Similarly, one may also not set the "magnetic current" equal to zero if one is to avoid a coincidence.

It is then very unrealistic to imagine that such a manifold of quantities should, in fact, arise. In consideration of the fact that we have only one vector potential, the appearance of two vectors – viz., S_i and M_i – cannot be understood. One can conjecture that one should set the one vector equal to zero and the resulting agreement could be used as the field equations for the vector potential. However, it does not seem possible to bring the Δ-equations of the vector potential into such a relation. One must then accept the total system, which is extended by the four equations for the vector potential, or else the three vectors S_i, M_i, Φ_i would have an intrinsic relationship to each other. Naturally, that seems entirely unbelievable.

Furthermore, we also have absolutely no reference point for understanding the singular correction terms that enter the equations as a result of the vector potential.

Finally, we would like to exhibit a peculiar complication that is based in the fact that the equations of the dual objects appear to be mixed with the non-dual ones.

The dual objects have the following tensor-analytic meaning: Along with the fundamental tensor g_{ik} , which is a tensor of second order, there is also a second fundamental tensor of n^{th} order (so it is of fourth order in a four-dimensional space) on any manifold, which has the following remarkable property: It is anti-symmetric in all of its indices. Thus, all components for which two indices are equal must vanish. The non-vanishing components will be defined as follows: The covariants are equal to:

$$\sqrt{g}\;\eta_{iklm}\,, \tag{20}$$

while the contravariant ones are equal to:

$$\frac{1}{\sqrt{g}}\;\eta^{iklm}\,, \tag{20′}$$

where η has the following meaning: $\eta = +\,1$ when the permutation $iklm$ of the four numbers 1 to 4 is even and $\eta = -\,1$ when the permutation is odd. If one defines the invariant:

$$\frac{1}{\sqrt{g}}\;u_\mu\,v_\nu\,w_\rho\,r_\sigma\,\eta^{\mu\nu\rho\sigma} \tag{21}$$

with the help of this tensor, from four vectors then one obtains the determinant of the four vectors in the denominator, and we would then like to refer to this tensor as the "determinant tensor."

The dual objects of tensor analysis are now exhibited precisely with the help of this singular tensor by ordinary multiplication. One can – e.g., from an anti-symmetric tensor F_{ik} of second order – construct the following new tensor:

$$\tilde{F}^{ik} = \frac{1}{2}\frac{1}{\sqrt{g}} F_{\mu\nu}\,\eta^{\mu\nu ik}. \tag{22}$$

This is precisely the "dual" tensor. In this way, one can, in any case, obviously make any anti-symmetric structure of order m correspond to an anti-symmetric dual structure of order $n - m$.

The determinant tensor is precisely a tensor like any other, except that it has the remarkable property that, by definition, a square root must be taken, from which, an indeterminacy of the sign emerges. One easily convinces oneself that under a reflection the determinant tensor behaves like an ordinary tensor under that transformation, except that the sign of the square root changes. In order to establish the sign of $\sqrt{g}$, one must proceed in such a way that one prescribes – say – the value +1 in all "right-handed" systems and –1 in all "left-handed" ones. There then exists no possibility of characterizing "right-handed coordinate systems" by invariant principles.

Thus, all that remains is the possibility of normalizing the sign of $\sqrt{g}$ consistently to be, say, – 1. Then, however, all objects that are constructed from the determinant tensor have the remarkable property that under its transformation the usual transformation formula are provided with the factor – 1 when one performs a transformation with negative determinant (reflection). (They are the so-called "axial" objects, in contrast to the "polar" ones.)

Permit me to explain this phenomenon, which is perhaps not generally known in the present context, by an example from three-dimensional tensor analysis. One can write the first system of Maxwell vacuum equations (time is now considered to be an ordinary scalar quantity) as follows:

$$\frac{1}{c}\frac{\partial E^i}{\partial t} = \frac{1}{\sqrt{g}}\frac{\partial H_\nu}{\partial x_\mu}\eta^{\mu\nu i} \tag{23}$$

(g is now the determinant of the spatial line element.) We can now proceed by establishing that the sign of $\sqrt{g}$ is either positive or negative according to whether we are dealing with a right-handed or left-handed system, respectively (which, however, as we pointed out, is not necessarily an invariant distinction, and this does not conform to the spirit of general covariance). E_i and H_i are then both ordinary vectors, and the change of sign under a reflection takes place inside the equation. Otherwise, we normalize the sign of $\sqrt{g}$ to + 1, so either H_i or E_i must be an "axial" vector. As is known, this complication vanishes in the four-dimensional context. There, one can avoid the dual object completely *, if one so desires. However, it is also when one preserves it that no

* A. Einstein 1916, see W. Pauli, *Relativitätstheorie* (Teubner 1921), pp. 631.

dual objects seem to be mixed with non-dual ones, so the indeterminacy of $\sqrt{g}$ does not arise.

In our system of equations (18), this is, however, the case, in fact. If we would like to establish an invariant standpoint then we would be forced to regard M_i as an "axial vector," while M is an "axial scalar." It is remarkable that we must, however, also broadly consider Φ_i to be an axial vector; this is thus excluded. If the less-likely case in which the vector potential is an axial scalar were chosen then it would also be the electric charge. However, since the velocity is certainly polar the charge must be an axial scalar. Now, the charge is multiplied by the vector potential in Φ_i . The change of sign then vanishes in this case, as well, and a polar vector results in all cases.

Thus, nothing still remains but the two-valued nature of $\sqrt{g}$ and the distinction between "right-handed" coordinate system and the "left-handed" ones, which implies a certain concession to general covariance.

It is of interest to remark that this complication does not emerge in the new Einstein geometry of "teleparallelism." There, the metric is already constructed from fundamental quantities quadratically, so the determinant of the fundamental quantities enters in place of the root of the determinant. The determinant tensor also behaves there exactly like an ordinary tensor under reflections, and the difference between polar and axial quantities becomes trivial. Therefore, the dual objects in this theory have a more natural character than they do in Riemannian geometry, in which the anti-symmetric element does not, by any means, represent something foreign in and of itself.

We have succeeded formally in bringing the Dirac equation into a form in which the demands of conventional tensor calculus are completely justified and suggests a purely field-theoretic description in itself. When the system of equations thus obtained is not satisfied at various points and singles out incomprehensible elements then that must perhaps not be understood to mean that the path taken here is completely misleading. The close relationship of the Dirac operator to Hamilton's quaternion operator, on the one hand, and the close relationship between this operator and four-dimensional tensor analysis, on the other, might then suggest that the connection that we discovered here must be more than just an odd coincidence. Perhaps the field-theoretic viewpoint that was assumed here, which only allows those operations that have some tensor-analytic sense, and also imparts a real interpretation to the imaginary quantities [*], implies heuristic Ansätze for a natural expansion of the Dirac theory, under which a transcription into a tensor-analytic system of equations comes about that proves to contain more intrinsic plausibility than the one here.

Berlin-Nikolassee, end of July, 1929.

[*] In the present representation, the *i* corresponds to the tail end of the tensor-analytic operation on dual objects.