Wave Equations in Curved Space-Time

S. DATTA MAJUMDAR
Department of Physics, University College of Science, Calcutta, India
(Received December 4, 1961)

The known wave equation for the electron in a gravitational field is generalized for all kinds of elementary particles allowed by the present theories. The usual conditions for the Lorentz invariance of a first order equation are shown to be sufficient for the generalization.

1. GEOMETRIC FOUNDATION

THE problem of formulating Dirac's equation within the framework of the general theory of relativity has been solved satisfactorily by a number of authors.[1–6] Brill and Wheeler[5] have set up radial equations for a Dirac particle in a metric with spherical symmetry and have discussed at length the types of solutions[7] obtained in the case of the electron and the neutrino. The thoroughness with which the problem has been investigated has left little scope for further work. The object of the present paper is only to show that the theory developed for the electron, with suitable modifications, applies to any elementary particle. The whole theory is presented here from a slightly different viewpoint, and the generality of the approach is clearly brought forth.

In the case of the electron the various treatments are physically equivalent and lead to the same equation:

$$\gamma^\mu\left(\frac{\partial}{\partial x^\mu}-\Gamma_\mu\right)\psi-ik\psi=0,$$

where the γ^μ satisfy the generalized anticommutation relations $[\gamma^\mu,\gamma^\nu]_+=2g^{\mu\nu}$. The matrix elements of γ^μ and Γ_μ are functions of coordinates. To understand the nature of this coordinate dependence we draw at every point of space-time four directions, or vierbein, which are mutually orthogonal but otherwise arbitrary. If these four directions are taken to be the axes of a Galilean frame of reference, then the curvilinear components of an elementary vector will be connected with its Galilean components by an equation of the type

$$dx^\mu=\lambda_h{}^\mu dx^h \tag{1}$$

[see Brill and Wheeler,[5] Eq. (6)]. $\lambda_h{}^\mu$ obviously behaves as a contravariant vector with respect to the index μ under general coordinate transformations and as a covariant vector with respect to the other index h under vierbein rotation.[8] Throughout this paper Greek indices will be used for world tensors and Latin indices for vierbein tensors. Thus, g_{ij} will denote the constant metric tensor of Minkowski space. With the help of g_{ij}, $g_{\mu\nu}$, and their associates, the indices of $\lambda_h{}^\mu$ may be raised or lowered. These elementary points have been discussed many times and we simply note down the important properties[9] of $\lambda_h{}^\mu$:

$$\lambda_h{}^\mu\lambda_\nu{}^h=\delta_\nu{}^\mu,\quad \lambda_h{}^\mu\lambda_\mu{}^k=\delta_h{}^k,$$
$$\lambda_\mu{}^h\lambda_{h\nu}=g_{\mu\nu},\quad \lambda_h{}^\mu\lambda_{k\mu}=g_{hk},$$
$$dx^h=\lambda_\mu{}^h dx^\mu.$$

Although the name is due to Einstein, this vierbein formalism was developed by Ricci long before the birth of the general theory of relativity and seems especially suited for introducing gravitation into the theory of elementary particles. The mathematical methods employed here will, however, be borrowed from the illuminating account of these researches given by Levi-Civita.[8] In the terminology adopted by him, $\lambda_h{}^\mu$ and $\lambda_{h\mu}$ are the parameters and moments of the orthogonal congruences of lines generated by the vierbein system. Let us now examine the effects on these quantities of an arbitrary rotation of the vierbein. Such a rotation corresponds to a local Lorentz transformation $dx^{i*}=L^i{}_a dx^a$. From $L^i{}_j$ we can construct $L_i{}^j$, L^{ij}, and L_{ij} formally by raising or lowering indices with the help of the Minkowski tensor. For finite transformations

$$L^i{}_aL_j{}^a=\delta_j{}^i=L^a{}_jL_a{}^i,$$

and for infinitesimal transformations

$$L^i{}_j=\delta^i{}_j+\omega^i{}_j,\quad 0=\omega^i{}_j+\omega_j{}^i=\omega_{ij}+\omega_{ji}=\omega^{ij}+\omega^{ji}.$$

Also,

$$\lambda_{h*}{}^\mu=L_h{}^a\lambda_a{}^\mu,\quad \lambda_h{}^\mu=L^a{}_h\lambda_{a*}{}^\mu.$$

In terms of the parameters and moments the Lorentz transformation is, therefore,

$$\lambda_\mu{}^{h*}\lambda_k{}^\mu=L^h{}_k. \tag{2}$$

Next, following the procedure adopted by Levi-Civita[9] for deriving Ricci's coefficients of rotation we carry the vectors $\lambda_h{}^\mu$ by parallel displacement from the point P to

[1] H. Tetrode, Z. Physik **50**, 336 (1928).
[2] J. A. Schouten, J. Math. and Phys. **10**, 239 (1930–31).
[3] V. Vock, Z. Physik **57**, 261 (1929).
[4] E. Schrödinger, Sitz. ber. preuss. Akad. Wiss. Physik math. Kl. **24**, 2105 (1932). The author had no opportunity of consulting this paper the journal not being available.
[5] D. R. Brill and J. A. Wheeler, Revs. Modern Phys. **29**, 465 (1957).
[6] O. Klein, Arch. Math. Astron. and Phys. **34**, 1 (1947).
[7] In this connection J. Callaway, Phys. Rev. **112**, 290 (1958), remarks that the radial equations have no power series solution about $r=0$.

[8] The word "rotation" will be used everywhere in the sense of Lorentz transformation.
[9] T. Levi-Civita, *The Absolute Differential Calculus* (Blackie and Son Limited, London, England, 1954), pp. 206, 261–286.

a neighboring point P_1 with coordinate differences δx^μ. We thus get a Galilean frame F' at P_1 which will not, in general, coincide with the frame F^* previously set up at P_1. By an infinitesimal rotation, however, the two frames F' and F^* can be brought into coincidence. If the parameters and moments pertaining to the frames F' and F^* are denoted by a prime and an asterisk, respectively, then the Lorentz transformation corresponding to this infinitesimal rotation is, by Eq. (2),

$$L^h{}_k = \lambda_\mu{}^{h*}\lambda_k{}^{\mu'} = \left(\lambda_\mu{}^h + \frac{\partial \lambda_\mu{}^h}{\partial x^\alpha}\delta x^\alpha\right)(\lambda_k{}^\mu - \{\beta\alpha,\mu\}\lambda_k{}^\beta \delta x^\alpha)$$

$$= \delta^h{}_k + \lambda^h{}_{\beta;\alpha}\lambda_k{}^\beta \delta x^\alpha,$$

or

$$\omega_{hk} = \lambda_{h\beta;\alpha}\lambda_k{}^\beta \delta x^\alpha. \qquad (3)$$

The quantities $p_{hk\mu} = \lambda_{h\beta;\mu}\lambda_k{}^\beta$ are antisymmetrical in h and k as expected. These are connected with Ricci's coefficients of rotation by the relations $\gamma_{hkl} = p_{hk\mu}\lambda_l{}^\mu$.

2. THE WAVE EQUATION IN GENERAL RELATIVITY

Having established the necessary mathematical formulas, we now proceed to show that the most general wave equation for elementary particles can be put into a form covariant under coordinate transformations and invariant under vierbein rotation. Nothing more is needed for a physical theory. The general *form* of the equation is unaffected by the commutation relations for the γ matrices, and from it Dirac's equation is obtained as a special case by adopting the relation $[\gamma^\mu,\gamma^\nu]_+ = 2g^{\mu\nu}$.

Whatever may be the special form of the theory a relativistic wave equation, as it is understood at present, must satisfy the following general conditions:

(i) It must be a first-order equation of the form

$$\gamma^i \frac{\partial \psi}{\partial x^i} - ik\psi = 0;$$

(ii) Under a Lorentz transformation, $x^{i*} = L^i{}_a x^a$, ψ must transform according to a representation of the (proper) Lorentz group, $\psi^* = S\psi$;

(iii) The matrix S must satisfy the relation $S^{-1}\gamma^i S = L^i{}_a \gamma^a$. These conditions must hold in all cases. Different theories of elementary particles are obtained by postulating different commutation relations for the γ matrices which, as stated before, leave the form of the equation in general relativity unaltered. The implications of the conditions (ii) and (iii) are seen more clearly by considering an infinitesimal Lorentz transformation, for which

$$S = 1 + \tfrac{1}{2}\omega_{hk} S^{hk}. \qquad (4)$$

Here, $S^{hk} = -S^{kh}$ are the infinitesimal operators of the Lorentz group, sometimes also called the nucleus of the group. They must satisfy Lie's integrability conditions,

$$[S^{hk}, S^{ij}] = g^{hj}S^{ki} + g^{ki}S^{hj} - g^{hi}S^{kj} - g^{kj}S^{hi}, \qquad (5)$$

a relation of cardinal importance in all discussions on the Lorentz group. Since the six quantities ω_{hk} are arbitrary we have, from condition (iii),

$$[\gamma^i, S^{hk}] = g^{ih}\gamma^k - g^{ik}\gamma^h. \qquad (6)$$

It is to be noted that the approach is perfectly general and rests only on the three conditions (i), (ii), and (iii). No particular assumption is made about the form of S^{hk}; nor is S^{hk} assumed to be irreducible under proper Lorentz transformations.

In carrying over all these considerations to general relativity we assume, with the previous authors, that ψ behaves as a multicomponent scalar under coordinate transformations and, therefore, $\partial\psi/\partial x^\mu$ behaves as a covariant vector. The contracted product $\gamma^\mu \partial\psi/\partial x^\mu$ will, therefore, be a scalar if the γ^μ's are constructed from the constant γ^i's of special relativity according to the prescription

$$\gamma^\mu = \lambda_h{}^\mu \gamma^h. \qquad (7)$$

When the vierbein is given an arbitrary rotation the γ^μ change to $\gamma^{\mu*} = L_h{}^a \lambda_a{}^\mu \gamma^h$. The invariance of the theory under vierbein rotation then demands that

$$\gamma^{\mu*} = S\gamma^\mu S^{-1}, \quad \psi^* = S\psi. \qquad (8)$$

The necessity of including an extra term in the wave equation now becomes apparent. Since the rotation varies arbitrarily from point to point, $\partial S/\partial x^\mu$ will not vanish and, therefore, $\partial\psi/\partial x^\mu$ will not have the correct transformation properties. The extra term $-\Gamma_\mu\psi$ should be such that the whole quantity $\partial\psi/\partial x^\mu - \Gamma_\mu\psi$ transforms like ψ under vierbein rotation. This quantity may legitimately be called the covariant derivative of ψ and denoted by the convenient symbol $\psi_{;\mu}$.

To derive an expression for $\psi_{;\mu}$ we assume that after parallel displacement from P to P_1 the components of ψ in the frame F' remain unaltered. The passage to the frame F^* by an infinitesimal Lorentz transformation will then change ψ into $\psi^* = (1 + \frac{1}{2}\omega_{hk}S^{hk})\psi$. Inserting the value of ω_{hk} from Eq. (3) we have

$$\delta\psi = \tfrac{1}{2}p_{hk\alpha}S^{hk}\delta x^\alpha \psi = \Gamma_\alpha \psi \delta x^\alpha \qquad (9)$$

as the equation of parallelism for ψ. Therefore,

$$\psi_{;\mu} = \frac{\partial\psi}{\partial x^\mu} - \tfrac{1}{2}p_{hk\mu}S^{hk}\psi = \frac{\partial\psi}{\partial x^\mu} - \Gamma_\mu\psi, \qquad (10)$$

and the generalized wave equation reads

$$\gamma^\mu \psi_{;\mu} - ik\psi = 0. \qquad (11)$$

Once the covariant derivative of ψ has been found out the main difficulty is over, and the concept can be immediately generalized, with the help of the known results of tensor calculus, to the case of a mathematical

object which transforms as a tensor under coordinate transformations and according to a representation of the Lorentz group under vierbein rotation. $\psi_{;\mu}$ itself is a mixed quantity[10] of this type, and its covariant derivative will be

$$\psi_{;\mu\nu}=\frac{\partial\psi_{;\mu}}{\partial x^\nu}-\Gamma_\nu\psi_{;\mu}-\{\mu\nu,\alpha\}\psi_{;\alpha}. \tag{12}$$

The addition of the last term gives $\psi_{;\mu\nu}$ the necessary tensor character.

It remains only to show that γ^μ transforms according to Eq. (8), which implies that

$$\omega_{hk}\lambda^{k\mu}\gamma^h=-\tfrac{1}{2}\omega_{hk}[\gamma^\mu,S^{hk}]. \tag{13}$$

Multiplying this by $\lambda_\mu{}^i$ we have

$$g^{ih}\gamma^k-g^{ik}\gamma^h=[\gamma^i,S^{hk}]$$

and this is identical with Eq. (6). Next, taking the value of ω_{hk} from Eq. (3) we get the important relation

$$[\Gamma_\alpha,\gamma^\mu]=\partial\gamma^\mu/\partial x^\alpha+\{\beta\alpha,\mu\}\gamma^\beta. \tag{14}$$

Thus, we see that the entire theory of elementary particles, as it stands today, fits nicely into the scheme of general relativity. We have definite rules for translating every item of the theory into the language of general relativity. For instance, the generalized commutation relations are obtained by using Eq. (7).

Although there can be no doubt that $\psi_{;\mu}$, as defined by Eq. (10), has the correct transformation properties, we subject it to an interesting test obtaining, thereby, the value of an expression arising in the calculation of the divergence of the energy momentum tensor. Let us carry ψ by parallel displacement round a parallelogram Ω formed by the elementary vectors δx^μ and $\delta' x^\mu$ drawn from a point P. By inserting the row and column indices, the equation of parallelism (9) can be written as

$$\delta\psi_r=\Gamma_{\mu|rs}\psi_s\delta x^\mu\simeq\bigg[(\Gamma_{\mu|rs})_p+\bigg\{\Gamma_{\mu|rk}\Gamma_{\nu|ks}+\frac{\partial\Gamma_{\mu|rs}}{\partial x^\nu}\bigg\}(x^\nu-x_p{}^\nu)\bigg]\psi_s\delta x^\mu. \tag{15}$$

By Stokes theorem the line integral of this Pfaffian round Ω will be equal to

$$D\psi_r=\bigg[\frac{\partial\Gamma_{\beta|rs}}{\partial x^\alpha}-\frac{\partial\Gamma_{\alpha|rs}}{\partial x^\beta}+\Gamma_{\beta|rk}\Gamma_{\alpha|ks}-\Gamma_{\alpha|rk}\Gamma_{\beta|ks}\bigg]\times\psi_s\delta x^\alpha\delta' x^\beta. \tag{16}$$

This formula is perfectly general and is independent of the nature and number of the three-index symbols $\Gamma_{\mu|rs}$. Erasing the row-column indices we obtain

$$D\psi=\bigg[\frac{\partial\Gamma_\beta}{\partial x^\alpha}-\frac{\partial\Gamma_\alpha}{\partial x^\beta}-[\Gamma_\alpha,\Gamma_\beta]\bigg]\psi\delta x^\alpha\delta' x^\beta$$

as the increment of ψ after the cyclic displacement. The same increment can also be calculated by carrying the Galilean frame at P by parallel displacement round Ω. Replacing ψ_r in Eq. (16) by $\lambda_h{}^\mu$ and the $\Gamma_{\mu|rs}$ by Christoffel's symbols we have

$$D\lambda_h{}^\mu=\{\nu\mu,\alpha\beta\}\lambda_h{}^\nu\delta x^\alpha\delta' x^\beta.$$

The Lorentz transformation from this frame to the original frame at P is given by

$$\omega_{hk}=-(\mu\nu,\alpha\beta)\lambda_h{}^\mu\lambda_k{}^\nu\delta x^\alpha\delta' x^\beta.$$

So, we must have

$$\frac{\partial\Gamma_\beta}{\partial x^\alpha}-\frac{\partial\Gamma_\alpha}{\partial x^\beta}-[\Gamma_\alpha,\Gamma_\beta]=-\tfrac{1}{2}(\mu\nu,\alpha\beta)\lambda_h{}^\mu\lambda_k{}^\nu S^{hk}. \tag{17}$$

The correctness of this result can be verified directly by using the relation (5). The matrix on the left-hand side operating on ψ gives also the difference of the second covariant derivatives of ψ. This difference vanishes in flat space-time.

We add a few remarks about the possibility of constructing a charge current vector and an energy momentum tensor from the wave equation (11). For the construction of these physically important quantities it is essential that there should exist a matrix η such that $\gamma^{i*}=\eta\gamma^i\eta^{-1}$ is the Hermitian conjugate of γ^i. Without loss of generality η can be taken to be Hermitian. If the Hermitian conjugate ψ^* is replaced by $\psi^\dagger=\psi^*\eta$, then the expressions for the charge-current and the energy-momentum assume the usual forms only if $S^{hk*}=-\eta S^{hk}\eta^{-1}$.

3. SIMPLIFICATION IN AN IMPORTANT SPECIAL CASE

In the case of Dirac and Duffin-Kemmer equations the relation between S^{hk} and γ^i is $S^{hk}=c[\gamma^h,\gamma^k]$, where c is a constant c-number. Whenever S^{hk} has this special form the elaborate procedure of constructing covariant derivatives becomes unnecessary and the generalized wave equation can be derived[11] very simply as follows.

The only thing required for the derivation is that $\partial S/\partial x^\mu$ arising after the rotation of the vierbein system should be canceled by some term coming from Γ_μ.

[10] In the case of $\lambda_h{}^\mu$, another mixed quantity, the particular representation S^{hk} is the infinitesimal Lorentz transformation itself. In the case of a product of several quantities S^{hk} will be a direct product.

[11] This is how the author verified the known result for the electron. The suggestive form (18) was then tested and found to be valid in the general case.

$\partial S/\partial x^\mu$ can be cast into the form

$$2\frac{\partial S}{\partial x^\mu}=\frac{\partial\omega_{hk}}{\partial x^\mu}c[\gamma^h,\gamma^k]=\frac{\partial\omega^h{}_k}{\partial x^\mu}\lambda_{h\alpha}\lambda_\beta{}^k c[\gamma^\alpha,\gamma^\beta]$$
$$=[(\omega^h{}_k\lambda_\beta{}^k)_{;\mu}-\omega^h{}_k\lambda_\beta{}^k{}_{;\mu}]\lambda_{h\alpha}S^{\alpha\beta}$$
$$=\Delta(\lambda^h{}_{\beta;\mu}\lambda_{h\alpha})S^{\alpha\beta}\quad\text{(where }\Delta\text{ denotes change after a Lorentz transformation)}$$
$$=\Delta(\lambda^h{}_{\beta;\mu}\lambda_{h\alpha})SS^{\alpha\beta}S^{-1}.$$

This shows that

$$\Gamma_\mu=\tfrac{1}{2}\lambda^h{}_{\beta;\mu}\lambda_{h\alpha}S^{\alpha\beta}=\tfrac{1}{2}p_{hk\mu}S^{hk}. \tag{18}$$

Evidently, one can add to the wave equation terms of the type

$$[A+B_\mu\gamma^\mu+C_{\mu\nu}\gamma^\mu\gamma^\nu+D_{\mu\nu\rho}\gamma^\mu\gamma^\nu\gamma^\rho+E_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma+\cdots]\psi,$$

where, A, B_μ, etc. are world tensors. In the case of Dirac's equation the series terminates at the quartic term and the tensors $C_{\mu\nu}$, $D_{\mu\nu\rho}$, $E_{\mu\nu\rho\sigma}$ may be taken to be antisymmetrical in all the indices. But, such additional terms, whatever may be their significance, may require additional terms in the current.[12]

4. NON-UNIQUENESS OF THE COVARIANT DERIVATIVE

In Riemannian geometry, Levi-Civita's notion of parallelism has a well-defined meaning. If a long narrow strip is cut out from a curved surface and placed without kinks on a plane then the Levi-Civita parallel become parallel in the ordinary sense. Any modification of the definition of parallelism in Riemannian space would, therefore, be quite artificial. But, in the case of of a multicomponent wave-function parallelism has no clear physical meaning, and we are at liberty to change the definition as we wish. Such modifications lead to additional terms in the wave equation which are to a large extent arbitrary. In carrying a vector from P to a neighboring point P_1 let us give it an extra rotation depending linearly on the coordinate differences. A convenient expression for this rotation can be obtained by using the tensor character of rotation matrices. Let

$$A^{\mu\prime}=L^\mu{}_\nu A^\nu$$

be a vector correspondence in the curvilinear coordinate system. From the invariance of the scalar product of two arbitrary vectors, it follows, as in the case of flat space, that

$$L^\alpha{}_\nu L_\alpha{}^\mu=\delta_\nu{}^\mu=L^\mu{}_\alpha L_\nu{}^\alpha,$$
$$0=\omega^\mu{}_\nu+\omega_\nu{}^\mu=\omega_{\mu\nu}+\omega_{\nu\mu}=\omega^{\mu\nu}+\omega^{\nu\mu}.$$

These relations must hold if the vector correspondence is to be a Lorentz transformation. An infinitesimal Lorentz transformation can, therefore, be written as

$$\omega^\mu{}_\nu=\epsilon(\varphi^\mu\chi_\nu-\varphi_\nu\chi^\mu) \tag{19}$$

in terms of two arbitrary world vectors φ^μ, χ^μ. A vector perpendicular to both φ^μ and χ^μ is left unaltered by this transformation. Since the scalar product of two vectors also remains unaltered, Eq. (19) defines a true Lorentz transformation in the plane of φ^μ, χ^μ. The simplest expression of this type linear in δx^μ is

$$\omega^\mu{}_\nu=\varphi^\mu\delta x_\nu-\varphi_\nu\delta x^\mu.$$

Combined with the Levi-Civita displacement this rotation gives rise to a kind of screw motion of the vectors $\lambda_h{}^\mu$. The Lorentz transformation from the resulting Galilean frame $\bar{F}'$ at P_1 to the vierbein frame F^* originally set up at P_1 will contain an extra term

$$[-\lambda_{h\alpha}\lambda_{k\beta}+\lambda_{k\alpha}\lambda_{h\beta}]\varphi^\beta\delta x^\alpha,$$

and the modified covariant derivative will be

$$\frac{\partial\psi}{\partial x^\mu}-[\tfrac{1}{2}p_{hk\mu}-\lambda_{h\mu}\lambda_{k\beta}\varphi^\beta]S^{hk}\psi.$$

It is perhaps desirable to demonstrate clearly that this unconventional term does satisfy the requirement of vierbein invariance. Let this term be denoted by $qS^{hk}\psi$. Then vierbein invariance requires that

$$(q+\Delta q)S^{hk}S-qSS^{hk}=\Delta qS^{hk}+\tfrac{1}{2}\omega_{ij}q[S^{hk},S^{ij}]$$

should vanish. Here, Δ denotes the change resulting from an infinitesimal Lorentz transformation ω_{ij}. By the formulas given in §1

$$\Delta qS^{hk}=\omega_{ij}\varphi^\beta[\lambda_\mu{}^j\lambda_{k\beta}-\lambda_\beta{}^j\lambda_{k\mu}]S^{ik}.$$

With the help of the relation (5) it is easily shown that $\frac{1}{2}\omega_{ij}q[S^{hk},S^{ij}]$ has the same value but the opposite sign. Thus, the requirement is fulfilled.

In the case of the electron $S^{hk}=\frac{1}{4}[\gamma^h,\gamma^k]$, and the additional term in the wave equation takes the familiar form $\frac{3}{2}\gamma^\mu\varphi_\mu\psi$. The electromagnetic potentials are thus introduced by modifying the definition of covariant differentiation. The usual procedure for introducing them is to define Γ_μ by Eq. (14) and then to show that the solution is determined only up to an additive multiple of the unit matrix. Both the methods utilize an arbitrariness in the formulation of the problem and are purely formal.

The above example certainly does not exhaust all possibilities. Other choices of the extra rotation $\omega^\mu{}_\nu$ lead to other kinds of terms in the covariant derivative. For instance, if[13]

$$\omega^\mu{}_\nu=(F^\mu{}_{\nu\rho}-F_\nu{}^\mu{}_\rho)\delta x^\rho,$$

then the additional term in the covariant derivative is $-F_{\alpha\beta\mu}\lambda_h{}^\alpha\lambda_k{}^\beta S^{hk}\psi$. Whatever may be the significance of such terms, we have at least succeeded in showing that the theory admits of a generalization.

[12] W. Pauli, Revs. Modern Phys. **13**, 203 (1941).

[13] This appears to be the most general form of $\omega_\nu{}^\mu$. The previous form $\omega_\nu{}^\mu=\phi^\mu\delta x_\nu-\phi_\nu\delta x^\mu$ is obtained by putting $F_{\mu\nu\rho}=\phi_\mu g_{\nu\rho}$.