ANNALES DE L'I. H. P., SECTION A

A. A. EL-KHOLY
R. U. SEXL
H. K. URBANTKE

On the so-called "Palatini method" of variation in covariant gravitational theories

Annales de l'I. H. P., section A, tome 18, nº 2 (1973), p. 121-135
<http://www.numdam.org/item?id=AIHPA_1973__18_2_121_0>

Ann. Inst. Henri Poincaré,
Vol. XVIII, n° 2, 1973, p. 121-136.

Section A :
Physique théorique

On the so-called " Palatini method " of variation in covariant gravitational theories

by

A. A. EL-KHOLY (*), R. U. SEXL and H. K. URBANTKE (**)

Institut für Theoretische Physik,
Universität Wien

ABSTRACT. — A review and discussion is given of what has been called the " Palatini method of variation " of the action integral in several covariant theories of gravitation. There are at least two versions of this " method " : One of them gives metric affinities in all cases, i. e. the (strong) principle of equivalence; the other one leads to deviations in some cases, among the effects of which there are causal anomalies. It is shown how, with spinorial affinities treated in the second way, gravity and electromagnetism can be unified.

1. INTRODUCTION

In Einstein's original version of the theory of gravitation [1] he gave a variational principle for the field equations, in which the Lagrange density

$$\mathcal{L} = \sqrt{-g}\, g^{ik}\, R_{ik} \tag{1}$$

is a function of the metric and its first and second derivatives (" purely metric theory ") : the affinities $\Gamma^i{}_{mn}$ from which the Ricci tensor

$$R_{ik} = \Gamma^j{}_{ij,k} - \Gamma^j{}_{ik,j} + \Gamma^j{}_{im}\,\Gamma^m{}_{jk} - \Gamma^j{}_{mj}\,\Gamma^m{}_{ik} \tag{2}$$

(*) Supported by grant from the Österreichisches Bundesministerium für Wissenschaft und Forschung.

(**) Supported by the « Fonds zur Förderung der wissenschaftlichen Forschung in Österreich ", Nr. 1534.

is constructed were considered to be the Christoffel symbols,

$$\Gamma^j{}_{mn} = \{ m^j n \} = \frac{1}{2} g^{ik} \left(- \frac{\partial g_{mn}}{\partial x^k} + \frac{\partial g_{kn}}{\partial x^m} + \frac{\partial g_{km}}{\partial x^n} \right). \tag{3}$$

The actual variation with respect to the metric was carried out in detail by Palatini [2] according to the scheme

$$0 = \frac{\delta \mathcal{L}}{\delta g_{ik}} = \frac{\partial \mathcal{L}}{\partial g_{ik}} + \frac{\partial \mathcal{L}}{\partial \{ m^j n \}} \frac{\delta \{ m^j n \}}{\delta g_{ik}}, \tag{4}$$

where $\delta \{ m^j n \}$ is calculated explicitly from the Christoffel relation (3). Later so-called " mixed affine-metric " theories were developed ([3], [4]) in which g_{ik}, $\Gamma^j{}_{mn}$ were considered as *independent* field variables in some suitable Lagrange density. It was noted that if (1) is treated in this way, the equations

$$0 = \frac{\delta \mathcal{L}}{\delta g_{ik}}, \qquad 0 = \frac{\delta \mathcal{L}}{\delta \Gamma^j{}_{mn}} \tag{5}$$

are equivalent to the system (3)-(4). In most of the literature (1) this observation is ascribed to Palatini, quoting [2], where it cannot be found; after this remark we shall, however, continue to call the separate independant variation of g_{ik} and $\Gamma^j{}_{mn}$ the " Palatini method ", until we are forced to make the distinctions which are part of the purpose of this article.

When an electromagnetic Lagrangian is added to (1), both kinds of variations still remain equivalent, if the affine connexion is assumed to be free of torsion [6]. We shall assume vanishing torsion (" symmetric affinities ") throughout this paper.

The Palatini method is considered to be more natural by Weyl, as he sees no reason to regard the relation between the affine and metric properties of the space as given. Besides, it has the following advantages :

(i) It allows for generalizations of General Relativity, since considering $\Gamma^j{}_{mn}$ to be independent of g_{ik} one can drop the symmetry condition on $\Gamma^j{}_{mn}$. In fact, the method was used by Einstein in an attempt to unify gravitation and electromagnetism by considering g_{ik}, $\Gamma^j{}_{mn}$ both to be non-symmetric [3].

(ii) It allows construction of theories invariant under the " Weyl gauge " [7],

$$g_{ik} \to \lambda(x)\, g_{ik}, \qquad \Gamma^j{}_{mn} \to \Gamma^j{}_{mn}. \tag{6}$$

(1) A notable exception is Pauli's book [5].

(iii) If g_{ik}, $\Gamma^{j}{}_{mn}$ are considered to be independent fields, the factor ordering problem, which occurs in attempts to quantize the theory, is somewhat simpler. Thus the Palatini method is more suitable for quantization.

Before one can use the " Palatini method ", however, the question has to be settled whether it is at all applicable in cases where the gravitational field is coupled to other fields other than the electromagnetic or scalar field, e. g., to a vector meson field. Weyl [6] has shown as early as 1948 that in the case of coupled gravitational and electron fields the ordinary procedure of varying g_{ik} alone, with (3) imposed (henceforth called " g-variation ") is no more equivalent to an independent variation of metric and affinity (henceforth called " P-variation "). His interpretation was that : " ... by the influence of matter a slight disaccordance between affine connection and metric (affine connection) is created ". Then he proposed a change in the Lagrangian leading to coincidence again. We point out, however that this coincidence, achieved through the introduction of the Weyl term in the Lagrangian, is merely a formal one, as the equation determining the affinity is not changed by this term.

The present work is a study of the use of the Palatini method in cases where the gravitational field is coupled to other fields. Section 2 gives a brief review which analyzes what may be meant by the Palatini method. Two possible versions are given : one has to distinguish between a " Palatini principle " (PP) and the " formal Palatini method of variation " (FP). The latter is simply a transformation of variables (a kind of Legendre transformation) and thus always equivalent to g-variation, although it may sometimes have some advantages. The PP, on the contrary, is in general not equivalent to g-variation, as shown in Section 3. Therefore it must lead to theories with new features (which is the reason for calling it a " principle "). If such a principle is adopted, one must study its consequences.

One of these consequences is, as we shall show in Section 4, the occurrence of causal anomalies. This will be illustrated by the simplest non-trivial example, the coupling of a vector field. On the basis of this effect one should reject the PP, although the numerical amounts for this and other deviations of the theories derived from a PP will be rather small.

The special cases in which PP and g-variation are equivalent are considered in Section 5. In particular, it is shown there that this is the case for all free compensation fields, in Utiyama's sense [8].

Finally, an application of the PP using spinors is given in which the spinorial affinities may be related to the electromagnetic field, thus leading to a unification of electromagnetism and gravity.

2. THE DIFFERENT VERSIONS OF THE PALATINI METHOD

It has often been pointed out ([9], [10], [11]) that the g-variation and " Palatini method " are analogous to the two procedures employed in deriving the electromagnetic field equations from a Lagrangian by either considering the 4-potential A_i as the only independent variable, while the field tensor is defined by

$$F_{ik} = 2\,A_{[i,k]} \tag{7}$$

or alternatively by considering A_i and F_{mn} as independent, in which case (7) follows from the variation of the Lagrangian. Many examples of such a twofold approach can be found in the literature; two of them are indicated in the Table.

TABLE

Usual Variation		Formal Palatini Method	
Lagrangian	Variables	Lagrangian	Variables
$L = -\frac{1}{2} A_{i,k}(A^{i,k} - A^{k,i})$	A_i	$\bar{L} = A_{i,k} F^{ik} + \frac{1}{4} F_{ik} F^{ik}$	A_i, F_{ik}
$L = \frac{1}{2}(\varphi_{,i}\,\varphi^{,i} - m^2 \varphi^2)$	φ	$\bar{L} = \frac{1}{2}\psi^i (2\,\varphi_{,i} - \psi_i) - m^2 \varphi^2$	φ, ψ_i

In fact, as pointed out by de Witt [12], neither the number of variables nor the form of the Lagrangian is basically relevant for the development of a given system. But these considerations together with a remark of Wheeler [13], that this second kind of variation is simply a transition from a Lagrange to Hamilton kind of formalism shows that there is an important difference between all these cases (e. g. of the Table) and the " Palatini method " described so far : for the latter the form of the Lagrangian was *not* changed. One could make this distinction more obvious by treating the pure gravitational case exactly in this Hamiltonian way, which in the point-mechanical case means to vary not the original Lagrangian $L(q, \dot{q})$ but the Lagrangian

$$\bar{L} = p\,\dot{q} - H(p, q)$$

(H being the Hamiltonian) with respect to q, p to obtain the equations of motion in the canonical form. The transition from variables q to q, p and from L to $\bar{L}$ gives the transition from the left to the right column in the Table. If one starts with the well-known first order form of the Lagrangian for gravity,

$$\mathfrak{G} = \sqrt{-g}\, g^{mp} \left(\Gamma_{m}{}^{r}{}_{q}\, \Gamma_{p}{}^{q}{}_{r} - \Gamma_{m}{}^{q}{}_{p}\, \Gamma_{q}{}^{r}{}_{r}\right) \tag{8}$$

with $\Gamma_{m}{}^{j}{}_{n}$ given by (3), the procedure just described must of course yield a system of equations equivalent to Einstein's. If one, however, treated (8) by the " Palatini method ", varying g_{ik} and $\Gamma^{j}{}_{mn}$ in (8) directly, one would not only get disagreement with Einstein's equations but also get non-covariant equations. This is so, because the difference between (8) — which is not a scalar density — and (1) is a pure divergence only if (3) is assumed from the beginning, but not if $\Gamma^{j}{}_{mn}$ are independent.

These considerations lead us towards making a distinction between two procedures which seem to have been called sometimes " Palatini method of variation " :

(i) The " formal Palatini method " is a mathematical procedure to treat a given theory by introducing new independent fields in such a way as to lower the order of the field equations and to construct a Lagrangian for the new form of the field equations. It proceeds in the same way as the transition from a Lagrangian to a Hamiltonian formalism.

(ii) The " Palatini Principle " is the following prescription to construct a Lagrangian and to obtain the field equations which define a theory of gravitation and its interaction with matter fields :

(1) form the Riemann and Ricci tensor for an arbitrary (torsionfree, as we shall assume) affine connection and contract with an arbitrary metric to obtain the Lagrange density of the free gravitational field;

(2) from the Lagrangian of the (interacting) matter fields in flat space obtain the Lagrangian describing also their gravitational interaction by replacing the flat metric by g_{ik} and ordinary derivatives by covariant ones with respect to the affine connection (as index transport and covariant derivation do not commute any more, this prescription is not quite unique);

(3) Vary the sum of both Lagrangians with respect to matter field variables, metric and affinities separately to obtain the field equations. The total Lagrangian (density) must be a scalar density or differ from a scalar density only by a pure divergence expression (the latter must be a divergence expression without assuming any relation between all field quantities, i. e. matter variables, metric, affinity).

In the following, we want to investigate some consequences of theories constructed according to the PP. We already mentioned that it is equivalent to ordinary General Relativity for pure gravitation. As the prescription (2) deviates from the ordinary (metric) " substitution principle " or " minimal coupling " [14] which guarantees that the weak principle of equivalence is satisfied, one could think of a violation of the latter. However, since in equivalence principle considerations one always deals with " test fields ", the deviations from the metric theory must be neglected in this respect. Other observable effects will be very (i. e. unmeasurably, under normal circumstances) small. Therefore we will, after exhibiting an example where deviations should occur, turn to a more basic matter : causal behavior.

3. PP TREATMENT OF A VECTOR FIELD

One of the simplest cases where the PP leads to a theory which differs from pure metric theory is the case of a vector field. For simplicity we take it to be massless and do not project out all spins except one, but start simply with a Lagrangian of the form $A^{r,s} A_{r,s}$. Its interaction with gravity is described by the Lagrangian (the coupling constant is absorbed into A_r or has been set $= 1$) :

$$\mathcal{L} = \sqrt{-g}\,(R + g^{tr} g^{ls} A_{r.s} A_{t.s}) \tag{9}$$

with $R = g^{ik} R_{ik}$, given by (2), and the covariant derivative

$$A_{r.s} = A_{r,s} - \Gamma^k{}_{rs} A_k \tag{10}$$

also formed with the affinities $\Gamma^j{}_{mn}$. The g-variation procedure would be to put $\Gamma^j{}_{mn} = \{ m^j n \}$ and to vary g_{ik}, A_l, which gives

$$R^{ik} - \frac{1}{2} g^{ik} R = \frac{1}{2} g^{ik} A_{r;s} A^{r;s} - A_{r;}{}^{i} A^{r;k} - A^{i}{}_{;r} A^{k;r} + (A^k A^{(i;l)} + A^i A^{(k;l)} - AA^{(i;k)})_{;l}, \tag{11}$$

$$A^{i;k}{}_{k} = 0, \tag{12}$$

where; indicates metric covariant derivatives. But now let us apply the PP ! The algebra becomes a little more complicated, because index transport by g_{ik} and covariant derivation are no more commuting operations. (For this reason we shall defer computational details to the appendix.) Hence (9) is no more a unique generalization of its special-relativistic form; but let us just consider (9). Varying g_{ik}, A_r, $\Gamma^j{}_{mn}$ gives

$$R^{ik} - \frac{1}{2} g^{ik} R = \frac{1}{2} g^{ik} A_{r.s} A_{t.}{}^{s} g^{rt} - g^{ti} g^{rk} A_{r.s} A_{t.s} - g^{tr} A_{r.}{}^{k} A_{t.}{}^{i}, \tag{13}$$

$$\Gamma^i{}_{jm} g^{jk} A_{k.}{}^{m} + \frac{1}{\sqrt{-g}} \partial_j (\sqrt{-g}\, g^{im} g^{jk} A_{m.k}) = 0, \tag{14}$$

$$
\begin{aligned}
(15) \quad & g^{rs}\left(\Gamma^{i}_{rs}\,\delta_j{}^k + \Gamma^{i}_{rs}\,\delta_j{}^i\right) + 2\left(g^{ik}\,\Gamma^{n}{}_{jm} - g^{rk}\,\Gamma^{i}{}_{rj} - g^{ri}\,\Gamma^{k}{}_{rj}\right) \\
& + 2\,\mathrm{A}_j\left(g^{mi}\,\mathrm{A}_{m;}{}^{k} + g^{mk}\,\mathrm{A}_{m.}{}^{i}\right) \\
& + \left(\ln\sqrt{-g}\right)_{,m}\left(g^{im}\,\delta_j{}^k + g^{km}\,\delta_j{}^i - 2\,g^{ik}\,\delta_j{}^m\right) \\
& \qquad + \left(\delta_j{}^k\,g^{im} + \delta_j{}^i\,g^{km} - 2\,g^{ik}\,\delta_j{}^m\right)_{.m} = 0,
\end{aligned}
$$

where R^{ik}, R are constructed from the Γ' s. Of course, these equations could be rewritten in tensor form, if one introduces the difference tensor $\Gamma^{j}{}_{mn} - \left\{ {j \atop m\,n} \right\}$. This difference tensor must be calculated from (15), which is done in the appendix; the result is complicated enough :

$$
\begin{aligned}
(16) \quad \Gamma^{m}{}_{ik} - \left\{ {m \atop i\;k} \right\} = {} & \beta\, g_{ik}\,\mathrm{A}^m \left\{ \frac{3-a}{2\,(1-a)}\,\mathrm{A}_{,r}{}^{r} + \frac{2\,a\,(3\,a-5)}{(1-a)\,(3+a)}\,\mathrm{A}^r\,\mathrm{A}^s\,\mathrm{A}_{,rs} \right\} \\
& + \frac{1}{1-a}\left\{ \mathrm{A}_i\,\mathrm{A}_k{}^{m} + \mathrm{A}_k\,\mathrm{A}_i{}^{m} - \mathrm{A}^m\,\mathrm{A}_{ik} \right\} \\
& + \frac{\mathrm{A}^j}{3+a}\left\{ \delta_k{}^m\,\mathrm{A}_{ji} + \delta_i{}^m\,\mathrm{A}_{jk} - 3\,g_{ik}\,\mathrm{A}_j{}^m \right\} \\
& + \beta\,\mathrm{A}_i\,\mathrm{A}_k\,\mathrm{A}^m \left\{ \frac{2a-3}{a-1}\,\mathrm{A}_{,r}{}^{r} + \frac{4\,(3\,a-5)}{(a-1)(a+3)}\,\mathrm{A}^r\,\mathrm{A}^s\,\mathrm{A}_{,rs} \right\} \\
& + 2\,\beta\,\mathrm{A}_{(i}\,\delta_{k)}{}^{m} \left\{ \frac{2\,a-3}{2(1-a)}\,\mathrm{A}_{,r}{}^{r} + \frac{2(a^2-2\,a-1)}{(a-1)\,(a+3)}\,\mathrm{A}^r\,\mathrm{A}^s\,\mathrm{A}_{,rs} \right\} \\
& + \frac{8}{(a-1)\,(a+3)}\,\mathrm{A}_i\,\mathrm{A}_k\,\mathrm{A}_j\,\mathrm{A}^{jm}
\end{aligned}
$$

with the abbreviations

$$
(17) \qquad a = \mathrm{A}^j\,\mathrm{A}_j, \qquad \beta = \frac{-1}{2\,a^2 - 2\,a - 3}, \qquad \mathrm{A}_{rl} = \mathrm{A}_{(r;\,l)}.
$$

It is obvious now that equation (14) with (16) does not reduce to equation (12), so that g-variation and PP are not equivalent in this case, as we wanted to demonstrate. We refrain from the actual insertion of (16) for obvious reasons.

4. PROPAGATION PROPERTIES OF THE VECTOR FIELD

As we stated in the Introduction, if one wants to take the PP seriously, one should be able to apply it for any coupled system and obtain reasonable results. The highly nonlinear equation that was obtained for the field A in Section 3 does not look too unreasonable, if one takes into account the weakness of the coupling; indeed, it reduces to (12) if terms quadratic in A are neglected. But there is one effect of the nonlinear terms which is more a matter of principle : they affect the propagation properties of signals, i. e. small discontinuities of A (or rather its second derivatives) travelling on a given back-ground g_{ik}, A_j.

To investigate this, we have to determine the condition for characteristic hypersurfaces for the field equations of A. According to standard methods, we first write down the terms in (14) which contain second derivatives. With the help of (16), we find (after extracting a factor 2-a which in general will be $\neq 0$) :

$$(18)\quad g^{im}\,\partial_m\,\partial_j\,A^i + PA^j\,A^m\,\partial_j\,\partial_m\,A^i + Q\,\partial^i\,\partial_j\,A^j + SA^j\,A^m\,\partial^i\,\partial_j\,A_m \\ + TA^i\,A^j\,A^m\,A^t\,\partial_j\,\partial_t\,A_m + SA^i\,A^j\,\partial_j\,\partial^m\,A_m + PA^i\,A^m\,g^{rs}\,\partial_r\,\partial_s\,A_m,$$

where P, Q, S, T are (rational) functions of $a = A^j\,A_j$ given in the Appendix. The characteristic hypersurfaces $\omega\,(x) = \text{const.}$ which belong to the field equation and some background g_{ik}, A_j have therefore to satisfy the following first order partial differential equation

$$(19)\quad \begin{cases} \det M_i^k = 0, \\ M_i^k = [\omega_j\,\omega^j + P\,(A_j\,\omega^j)^2]\,\delta_i^k + Q\,\omega_i\,\omega^k \\ \qquad + SA_j\,\omega^j\,(\omega_i\,A^k + A_i\,\omega^k) + [T\,(A_j\,\omega^j)^2 + P\,\omega_j\,\omega^j]\,A_i\,A^k. \end{cases}$$

We evaluate the determinant as the product of the four eigenvalues λ_1, λ_2, λ_3, λ_4 of M_i^k. First we observe that any vector orthogonal to ω^k, A^k is an eigenvector of M_i^k with eigenvalue $\omega_j\,\omega^j + P\,(A_j\,\omega^j)^2$. Since the vectors orthogonal to ω^k, A^k in general form a 2-dimensional space, our eigenvalue is twofold, hence

$$(20)\quad \lambda_1 = \lambda_2 = \omega_j\,\omega^j + P\,(A_j\,\omega^j)^2 = (g^{jl} + PA^j\,A^l)\,\omega_j\,\omega_l.$$

To find the other eigenvalues we must look for eigenvectors ξ_k,

$$(21)\quad M_i^k\,\xi_k = \lambda\xi_i$$

which are not simultaneously orthogonal to ω^k, A^k, i. e. $\omega^k\,\xi_k$ and $A^k\,\xi_k$ must not vanish simultaneously. Contracting (21) with ω^i, A^i we get two linear homogeneous equations for these two scalar products, if the explicit expression (19) is used. Putting the 2×2 determinant of this system equal to zero, we obtain a quadratic equation determining the other two eigenvalues. Their product is given by the λ-free term of this equation, i. e.

$$(22)\quad \lambda_3\,\lambda_4 = \begin{vmatrix} (1+Q)\,\omega_j\,\omega^j & [(P+S)\,\omega_j\,\omega^j \\ \quad + (P+S)\,(A_j\,\omega^j)^2 & \quad + T\,(A_j\,\omega^j)^2]\,A_j\,\omega^j \\ (Q + a\,S)\,A_j\,\omega^j & [(1 + a\,P)\,\omega_j\,\omega^j \\ & \quad + (P + S + a\,T)\,(A_j\,\omega^j)^2] \end{vmatrix}.$$

(22) is a binary quadratic form of the variables $\omega_j \omega^j$, $(A_j \omega^j)^2$ and may be factored as

$$(23) \quad \lambda_3 \lambda_4 = (1 + Q)(1 + a P)[\omega_j \omega^j + P' (A_j \omega^j)^2][\omega_j \omega^j + P'' (A_j \omega^j)^2],$$

where P', P'' are hereby defined new functions of $A_j A^j = a$ (they could even be conjugate-complex).

Collecting these results, we see that the local characteristic cone $\lambda_1 \lambda_2 \lambda_3 \lambda_4 = 0$ is split into three sheets, given by

$$(24\,a) \qquad (g^{jl} + P\ A^j A^l)\,\omega_j\,\omega_l = 0,$$

$$(24\,b) \qquad (g^{jl} + P'\,A^j A^l)\,\omega_j\,\omega_l = 0,$$

$$(24\,c) \qquad (g^{jl} + P''\,A^j A^l)\,\omega_j\,\omega_l = 0.$$

On the other hand, the characteristic cone for the gravitational field, and also for an additional scalar or electromagnetic field, has the usual equation (light cone) :

$$(24\,d) \qquad g^{jl}\,\omega_j\,\omega_l = 0.$$

Hence, against a non-vanishing background A^k, we have to expect several kinds of causal anomalies for the propagation of small discontinuities of A, which depend on the background and on the polarization of the discontinuity amplitude with respect to background and orientation of the initial discontinuity surface. There is one harmless, although interesting, case, in which all three cones (24 *a*, *b*, *c*) are normal-hyperbolic and inside the light cone (24 *d*); here one has to expect (locally, so that non-linearity plays not much role) phenomena like in crystal optics. In other cases one gets violation of causality or no propagation at all for some polarizations, etc., depending on the signature (elliptic, ultra-hyperbolic) of the cones and, in the normal-hyperbolic case, their position relative to the light cone.

More appropriate than to study all these possibilities in detail it would be in our opinion to exclude all of them by simply rejecting the PP as general principle. In cases of connections with torsion (2) this means that one should always require the connection to be metric and apply the PP only with this constraint.

5. ON THE CHOICE OF LAGRANGIANS

From the results of the previous sections as well as from the work of Weyl on the Einstein-Dirac fields it is seen that in general g-variation and PP will be inequivalent. One might, of course, object that the

(2) Connections with torsion recently received some interest, in particular by Trautman and his co-workers [15].

Lagrangian of a physical system is not " general ", and that for all correctly chosen physical Lagrangians one would get equivalence or at least not nonsensical features. As an example, the requirement for the vector field to carry pure spin 1 would force us to use the Lagrangian for the Proca or Maxwell field, where the affinities drop out when written covariantly in flat space, so that g-variation and PP are indeed equivalent in this case.

However, the correct Lagrangian of the real world has not yet been found. What one can try is to make general statements about the types of Lagrangians where one will get equivalence, and where not. One can classify the Lagrangians which are used in physics according to the symmetry groups involved; but we do not know all symmetry groups [16]. For theories with gauge groups one can make a general statement : for generalized Yang-Mills fields (or compensation fields in Utiyama's sense [8]) interacting only with gravitation, PP and g-variation are equivalent. This is so, because the Yang-Mills Lagrangians are (space-time and internal) scalars constructed from the internal curvature tensor which does not involve the space-time affinity.

With these general remarks in mind, we turn to a spinorial version of the PP which was discussed by Jamiołkowski [17]. Here the metric is defined with the help of the contravariant Pauli matrices

$$g^{j\dot{A}B}(x) \qquad (A,\ B = 1,\ 2) \tag{25}$$

i. e. an object transforming as a contravariant vector under coordinate transformations and as a contravariant Hermitian density under spinor transformations (*see* [17] for all details). Covariant Pauli matrices are defined by

$$\frac{1}{2}\, g_{j\dot{C}D}\, g^{j\dot{A}B} = \delta^{\dot{A}}{}_{\dot{C}}\, \delta^{B}{}_{D} \quad \leftrightarrow \quad g_{j\dot{A}B}\, g^{m\dot{A}B} = \delta^{m}{}_{j}. \tag{26}$$

From these, the real (symmetric) metric tensor is defined as

$$g_{kj} = \frac{1}{2}\, \varepsilon^{\dot{R}\dot{S}}\, \varepsilon^{PQ}\, g_{k\dot{R}P}\, g_{j\dot{S}Q}, \tag{27}$$

ε^{PQ} being the spinor metric. Another object defined from these quantities is the spin tensor

$$S^{kjAB} = \frac{1}{2}\, \varepsilon_{\dot{R}\dot{S}} \left(g^{k\dot{R}A}\, g^{j\dot{S}B} + g^{k\dot{R}B}\, g^{j\dot{S}A} \right) \tag{28}$$

satisfying

$$S^{kjAB} = S^{kj(AB)} = S^{[kj]AB}. \tag{29}$$

The tensorial and spinorial affinities $\Gamma^{j}{}_{mk}(x) = \Gamma^{j}{}_{km}(x)$, $\Gamma^{A}{}_{Bj}(x)$ are introduced independently from each other and from the $g_{j\dot{A}B}$. They describe how the operator of covariant differentiation, ∇_l, acts on tensors and spinors. The tensorial and spinorial curvature tensors are constructed from them in the usual way :

$$\text{(30)} \qquad \begin{cases} R^j{}_{mrs} = \Gamma^j{}_{ms,r} - \Gamma^j{}_{mr,s} + \Gamma^l{}_{ms}\Gamma^j{}_{lr} - \Gamma^l{}_{mr}\Gamma^j{}_{ls}, \\ B^J{}_{Mrs} = \Gamma^J{}_{Ms,r} - \Gamma^J{}_{Mr,s} + \Gamma^N{}_{Ms}\Gamma^J{}_{Nr} - \Gamma^N{}_{Mr}\Gamma^J{}_{Ns}. \end{cases}$$

The relation between metric and affine structure is now to be obtained "dynamically" from a variation principle. Jamiołkowski first considered the Lagrange density

$$\text{(31)} \qquad \sqrt{-g}\,(g^{jm} R_{jm} + S^{jm\,J}{}_M B^M{}_{Jjm} + \text{c. c.}),$$

but found that upon variation with respect to $g_{j\dot{A}B}$, Γ, $\Gamma^j{}_{mk}\Gamma^A{}_{Bj}$ one does not obtain enough independent equations to calculate $\Gamma^A{}_{Bj}$. Trying to overcome this difficulty, he introduced an additional antisymmetric, purely imaginary tensor field f_{jm}, for which he added the terms

$$\text{(32)} \qquad \varkappa\sqrt{-g}\,(f^{jm} B^N{}_{Njm} + \text{c. c.} + f^{jm} f_{jm})$$

to the above Lagrangian. From the resulting variational equations one derives

$$\text{(33)} \qquad \begin{cases} \Gamma^j{}_{km} = \left\{ {j \atop k\,m} \right\}, \qquad R_{jm} - \dfrac{1}{2} g_{jm} R = -8\pi\varkappa T_{jm}, \\ T_{jm} = \dfrac{-1}{4\pi}\left(\dfrac{1}{4} g_{jm} f^{rs} f_{rs} - f^r{}_j f_{rm}\right), \end{cases}$$

and $-if_{jm}$ satisfies the full Maxwell vacuum equations, which enables one to identify it with the electromagnetic field.

There are now two objections to be made : (*a*) The system of variational equations is still not enough to completely determine $\Gamma^A{}_{Bj}$, due to the assumption that f_{jm} is purely imaginary. The latter, however, is essential in order to have f_{jm} satisfying Maxwell's equations. (*b*) The introduction of a new field to obtain the relation between already existing fields is logically unsatisfactory. Enough relations of this kind should exist in the absence of this field, and it should even be expected that the new field will alter them in some cases.

These objections and difficulties can be overcome. Instead of f_{jm} one simply takes $B^N{}_{Njm}$, which is constructed from the already existing structures, and adds instead of (32) an Utiyama type term

$$\text{(34)} \qquad \varkappa\sqrt{-g}\,B^M{}_M{}^{jm} B^N{}_{Njm} + \text{c. c.}$$

The variations

(35 *a*) $\frac{\delta}{\delta\Gamma^j{}_{km}} \ldots = 0$ give 40 informations,

(35 *b*) $\frac{\delta}{\delta\Gamma^{\mathrm{A}}{}_{\mathrm{B}j}} \ldots = 0$ » 32 »

(35 *c*) $\frac{\delta}{\delta g^{j\,\dot{\mathrm{A}}\mathrm{B}}} \ldots = 0$ » 16 »

From (35 *a*) one concludes

(36 *a*) $$\nabla_j\, g_{mk} = 0 \;\rightarrow\; \Gamma^j{}_{mn} = \left\{ {j \atop m\, n} \right\}.$$

From (36 *a*) and (35 *b*) one gets, similar to [17] :

(36 *b*) $$\nabla_j\, \mathrm{B}^{\mathrm{N}}{}_{\mathrm{N}}{}^{lm} = 0 \quad \text{and} \quad \nabla_j\, \mathrm{S}^{jm\mathrm{AB}} = 0.$$

From these equations one can conclude, by arguments similar to corresponding calculations in [18] that $\mathrm{B}^{\mathrm{N}}{}_{\mathrm{N}jm}$ can be made purely imaginary, the quantity

(37) $$\mathrm{F}_{jm} = i\, \mathrm{B}^{\mathrm{N}}{}_{\mathrm{N}jm}$$

satisfying

(38) $$\nabla_j\, \mathrm{F}^{jm} = 0 = \mathrm{F}_{[jm,k]}$$

and, by (35 *c*),

(36 *c*) $$\mathrm{R}_{jm} - \frac{1}{2} g_{jm}\, \mathrm{R} = -\, 8\, \pi \varkappa\, \mathrm{T}_{jm},$$

(39) $$\mathrm{T}_{jm} = \frac{1}{4\,\pi} \left(\frac{1}{4} g_{jm}\, \mathrm{F}^{rs}\, \mathrm{F}_{rs} - \mathrm{F}^r{}_j\, \mathrm{F}_{rm} \right).$$

(38), (39) show that F_{jm} can be identified with the electromagnetic field. As $\mathrm{B}^{\mathrm{N}}{}_{\mathrm{N}jm}$ is geometric in character, we have thus obtained a kind of " unification " ([18]; note that $\mathrm{B}^{\mathrm{N}}{}_{\mathrm{N}jm}$ was not *assumed* to be purely imaginary from the beginning, so that (36 *b*) are enough equations to determine the $\Gamma^{\mathrm{A}}{}_{\mathrm{B}j}$).

APPENDIX

It is our purpose here to solve equation (15) for $\Gamma^j{}_{mn}$ in terms of A^j, g_{ik} and their derivatives. We introduce the tensor

(A.1) $$\mathrm{B}^m{}_{ik} = \Gamma^m{}_{ik} - \left\{ {m \atop i\;\; k} \right\} = \mathrm{B}^m{}_{ki}$$

and rewrite (15) as

$$g^{rs} B^{(i}{}_{rs}\, \delta_l{}^{k)} + g^{ik} B^m{}_{lm} - 2\, g^{r(i} B^{k)}{}_{rl} + 2\, A_l\, A_{m.}{}^{(i}\, g^{k)m} = 0 \tag{A.2}$$

[the equivalence with (15) becomes immediate if a coordinate system is used where the $\left\{ {m \atop i\ k} \right\}$ vanish at the point under consideration]. (A.2) can be contracted in two ways : first put $i = l$ and sum, to obtain

$$g^{rs} B^k{}_{rs} = -\frac{2}{3} C^k, \tag{A.3}$$

where

$$C^k = 2\, A_l\, A_{m.}{}^{(l}\, g^{k)m}; \tag{A.4}$$

then contract with g_{ik} to get

$$B^m{}_{lm} = \frac{1}{3} C_l - A_l\, A_{m.}{}^m. \tag{A.5}$$

(A.3, 5) are now substituted back into (A.2). Lowering all free indices we have

$$\frac{1}{3}(g_{kl}\, C_i + g_{li}\, C_k - g_{ik}\, C_l) + B_{ikl} + B_{kil} + g_{ik}\, A_l\, A_{m.}{}^m - 2\, A_l\, A_{(i.k)} = 0. \tag{A.6}$$

By taking cyclic permutations of (ikl), adding two of the resulting equations and substracting the third one, the following formal solution for B_{lik} results :

$$B_{lik} = \frac{1}{3} g_{l(i}\, C_{k)} - \frac{1}{2} g_{ik}\, C_l - A_{m.}{}^m \left(g_{l(i} A_{k)} - \frac{1}{2} g_{ik}\, A_l \right) + A_i\, A_{(k.l)} + A_k\, A_{(i.l)} - A_l\, A_{(i.k)}. \tag{A.7}$$

This does not yet give the $\Gamma^m{}_{ik}$, because these quantities appear also on the right hand side in the expressions $A_{m.n}$; we have

$$A_{i.k} = A_{i;k} - B^m{}_{ik}\, A_m. \tag{A.8}$$

Now we try to calculate $A_{i.k}$ and then obtain B_{lik} from (A.7). First we write

$$A_{i.k} = A_{[i.k]} + A_{(i.k)} = A_{[i;k]} + A_{(i.k)}, \tag{A.9}$$

where $A_{(i.k)}$ is implicitly defined if (A.7) is used :

$$\text{(A.10)}\quad A_{(i.k)} = A_{(i;k)} - B_{lik}A^l = A_{(i;k)} - \frac{4}{3}A_{(i}C_{k)} + \frac{1}{2}g_{ik}A^l C_l + A_{m.}{}^m\left(A_i A_k - \frac{1}{2}g_{ik}\,a\right) + a\,A_{(i.k)},$$

where

$$\text{(A.11)}\quad \begin{cases} C_k = 2\,A^l A_{m.(l}\,\delta_{k)}{}^m = 2\,A^l A_{(k.l)}, \\ a = A^l A. \end{cases}$$

(A.9, 10) lead to the following still implicit expression for $A_{i.k}$:

$$\text{(A.12)}\quad A_{i.k} = A_{[i;k]} + \frac{1}{1-a}\left[A_{(i;k)} + \left(A^l A^m A_{l.m} - \frac{1}{2}A_{m.}{}^m\,a\right)g_{ik} - \frac{4}{3}A_{(i}C_{k)} + A_{m.}{}^m A_i A_k\right].$$

Transvecting (A.12) with g^{ik}, $A^i A^k$, we obtain two linear scalar equations for the unknown scalars $A^l A^m A_{l.m}$, $A_{m.}{}^m$ which appear on the right hand side of (A.12). Their solution is

$$\text{(A.13)}\quad \begin{cases} A^l A^m A_{l.m} = \beta\left(\frac{3}{2}a^2 A^m{}_m + 3\,A^l A^m A_{lm}\right), \\ A_{m.}{}^m = \beta\left((2\,a+3)\,A^m{}_m + 4\,A^l A^m A_{lm}\right) \end{cases}$$

where the abbreviations (17) have been used.

Transvecting (A.12) with A^k only, we get an equation for C_i whose solution is

$$\text{(A.14)}\quad C_i = \frac{+3}{a+3}[\beta(2(2\,a-1)\,A^m A^l A_{ml} + a(a+3)\,A^m)\,A_i + 2\,A^m A_{im}].$$

Insertion of (A.13, 14) into (A.12) gives the desired explicit expression for $A_{i.k}$:

$$\text{(A.15)}\quad A_{i.k} = A_{[i;k]} + \frac{1}{1-a}A_{ik} + \frac{\beta}{1-a}\left[\frac{a}{2}(a-3)\,A^m{}_m - (2\,a-3)\,A^m A^j A_{mj}\right]g_{ik} - \frac{\beta}{(1-a)(a+3)}[(2\,a+3)(a+3)\,A^m{}_m + 4\,(3\,a-5)\,A^m A^j A_{mj}]\,A_i A_k - \frac{8}{(1-a)(a+3)}A^m A_{m(i}A_{k)},$$

whose further insertion into (A.7) gives B_{lik} in explicit form, i. e. (16). Thus the relation (A.1) between the " dynamical " connection $\Gamma^j{}_{mn}$ and the Christoffel connection is determined.

The terms in the field equation (14) which contain second derivatives after the substitution of (A.15) become (18), with the coefficients given by

$$(\text{A.16}) \quad \left\{ \begin{aligned} & \mathrm{P} = \frac{4}{(a-2)(a+3)}, \qquad \mathrm{Q} = \beta \frac{a^2(2a-3)}{(a-2)}, \\ & \mathrm{S} = -2\beta \frac{(a-3)(2a-1)}{(a-2)(a+3)}, \\ & \mathrm{T} = \beta \frac{8(3a-5)}{(a-2)(a+3)}. \end{aligned} \right.$$

REFERENCES

1] A. EINSTEIN, *Sitz. Ber. d. Preuss. Akad. Phys. Math. Kl.*, 1915, p. 844 et 1916, p. 1111.
[2] A. PALATINI, *R. C. Circ. Mat. Palermo*, vol. 43, 1919, p. 203.
[3] A. EINSTEIN, *Sitz. Ber. d. Preuss. Akad. Phys. Math. Kl.*, 1925, p. 414.
[4] E. SCHRÖDINGER, *Space-Time Structure*, Cambridge, 1950.
[5] W. PAULI, *The Theory of Relativity*, Pergamon, 1958.
[6] H. WEYL, *Phys. Rev.*, vol. 77, 1950, p. 699.
[7] G. STEPHENSON, *Nuovo Cimento*, vol. 9, 1958, p. 263.
[8] R. UTIYAMA, *Phys. Rev.*, vol. 101, 1956, p. 1597.
[9] J. L. ANDERSON, *Principles of Relativity Physics*, Academic Press, 1967.
[10] R. ARNOWITT et S. DESER, *Phys. Rev.*, vol. 113, 1959, p. 745.
[11] R. ARNOWITT, S. DESER et C. MISNER, *Phys. Rev.*, vol. 116, 1959, p. 1322 et vol. 117, 1960, p. 1595. Article in : *Gravitation, Introduction to Current Research*, ed. L. Witten, Wiley, 1962.
[12] B. S. DE WITT, Article in : *Relativity, Groups and Topology*, ed. de Witt and de Witt, Gordon and Breach, 1964.
[13] J. A. WHEELER.
[14] R. U. SEXL, *Fortschr. Phys.*, vol., 15, 1967, p. 269.
[15] A. TRAUTMAN, *Bull. Acad. Polon. Sci. Ser. Math. Astr. Phys.*, vol. 20, 1972, p. 185 et vol. 20, 1972, p. 503.
[16] E. P. WIGNER, *Physics Today*, vol. 17, March 1964.
[17] A. JAMIOLKOWSKI, *Reports on Math. Phys.*, vol. 2, 1971, p. 1.
[18] W. BADE et H. JEHLE, *Rev. Mod. Phys.*, vol. 25, 1953, p. 714.

(*Manuscrit reçu le* 8 *décembre* 1972.)