Quadratic Lagrangians and Palatini's device

1979 J. Phys. A: Math. Gen. 12 1229

(http://iopscience.iop.org/0305-4470/12/8/017)

J. Phys. A: Math. Gen., Vol. 12, No. 8, 1979. Printed in Great Britain

Quadratic Lagrangians and Palatini's device

H A Buchdahl

Department of Theoretical Physics, Faculty of Science, Australian National University, Canberra 2600, Australia

Received 11 October 1978

Abstract. This paper deals with the mutual inequivalence of g-variation and P-variation of a given action S, the components of linear connection $\Gamma^m{}_{kl}$ being required to be symmetric. Under g-variation S is required to be stationary with respect to variations of the metric tensor g_{ij}, the $\Gamma^m{}_{kl}$ being taken to be Christoffel symbols from the outset, whereas under P-variations the g_{ij} and $\Gamma^m{}_{kl}$ are initially regarded as mutually independent and S is required to be stationary with respect to independent variations of these quantities. The discussion is illustrated at length by examples in which the Lagrangian of S is one or another of a set of homogeneous or inhomogeneous quadratic invariants of the Riemann tensor.

1. Introduction

The field equations of relativistic theories of gravitation are often taken to be generated by a variational principle, i.e. they express the stationarity of some action $S := \int L(-g)^{1/2}\,dx$ under variations, with fixed boundary values, of the components g_{ij} of the metric tensor and their first derivatives. In particular, in Einstein's theory, which operates within the framework of a 4-dimensional normal-hyperbolic Riemann space V_4, the Lagrangian L is taken to be the scalar curvature R, at any rate to within an additive constant -2λ. In this case the variation of the g_{ij} brings with it the variation of the components $\Gamma^k{}_{ij}$ of the linear connection, since these are Christoffel symbols from the outset. In other words, the only quantities subject to independent variation are the g_{ij}; and I then speak of g-variation of S.

For formal reasons the procedure just described has sometimes been replaced by what is commonly referred to as 'Palatini's device', although Palatini was not responsible for it. In this process, here called P-variation, one regards the g_{ij} and the $\Gamma^l{}_{mn}$ as initially quite unrelated functions. (The $\Gamma^l{}_{mn}$ are here taken to be symmetric.) One then requires S to be stationary with respect to arbitrary independent variations of g_{ij} and $\Gamma^l{}_{mn}$ which vanish on the boundary. Thus, if $w := (-g)^{1/2}$ and $\mathfrak{X} := wX$ for every X, one has in the first place the generic identity

$$\delta\int \mathfrak{L}\,dx =: \int (\mathfrak{A}^{ij}\delta g_{ij} + \mathfrak{B}_k{}^{ij}\delta\Gamma^k{}_{ij})\,dx. \tag{1.1}$$

Consequently the 'P-equations' are

$$\mathfrak{A}^{ij} = 0, \tag{1.2}$$

$$\mathfrak{B}_k{}^{ij} = 0, \tag{1.3}$$

whereas the 'g-equations' are

$$(\mathfrak{B}^{jik} - \tfrac{1}{2}\mathfrak{B}^{kij})_{:k} - \mathfrak{A}^{ij} = 0, \tag{1.4}$$

where subscripts following a colon denote covariant differentiation with respect to the Christoffel symbols

$$\mathring{\Gamma}^{k}{}_{ij} := \tfrac{1}{2}g^{kl}(g_{li,j} + g_{lj,i} - g_{ij,l}). \tag{1.5}$$

One must not be deceived by superficial appearances: it is not the case that a given solution of the P-equations will necessarily satisfy the g-equations. If

$$\mathfrak{Z}_l{}^{ijk} := \partial\mathfrak{L}/\partial R^l{}_{ijk} \tag{1.6}$$

(symmetries of $R^l{}_{ijk}$ being ignored) and

$$\mathfrak{Y}_k{}^{ijl} := \tfrac{1}{2}(\mathfrak{Z}_k{}^{ijl} - \mathfrak{Z}_k{}^{ilj} + \mathfrak{Z}_k{}^{jil} - \mathfrak{Z}_k{}^{jli}), \tag{1.7}$$

then in (1.4)

$$\mathfrak{B}_k{}^{ij} = \mathfrak{Y}_k{}^{ijl}{}_{:l}, \tag{1.8}$$

whereas in (1.3)

$$\mathfrak{B}_k{}^{ij} = \mathfrak{Y}_k{}^{ijl}{}_{;l}, \tag{1.9}$$

subscripts following a semicolon denoting covariant differentiation with respect to $\Gamma^k{}_{ij}$; but the connections $\Gamma^k{}_{ij}$ and $\mathring{\Gamma}^k{}_{ij}$ need not, and in general will not, be the same. In short, the processes of g-variation and P-variation are in general inequivalent.

One gains the impression that the conclusion just stated is frequently not borne in mind when in various contexts P-variations are contemplated in place of g-variations—apparently as a mere matter of convenience. This may have come about because when $L = R - 2\lambda$ the two methods do happen to be equivalent. At any rate, the object of this paper is to examine P-variations in the context of quadratic Lagrangians (defined in § 2), supplementing a previous discussion (Buchdahl 1960) of difficulties associated with P-variations in general. The results obtained reinforce my belief that the uncritical use of Palatini's device, regarded merely as a convenient formal replacement for g-variations, is unjustified, for even where no internal inconsistencies arise one is confronted with the following position: a given Lagrangian (in general) gives rise to two distinct theories, one generated by g-variations and one generated by P-variations. The former operate from the outset in a Riemann space V_4, whereas the latter operate in a space A_4^* with symmetric linear connection on which an independent covariant tensor field of valence 2 is defined. In particular this implies that in the context of P-variations one is confronted with the presence of tensors such as $R_{[ij]}$, whereas under g-variations their absence is guaranteed by the Riemannian character of the connection. In short, a given theory will be generated by a specific Lagrangian L *together with* an explicit prescription that S be stationary *either* with respect to g-variations *or* with respect to P-variations.

2. Quadratic Lagrangians

L is said to be a 'quadratic Lagrangian' if it is a (homogeneous) quadratic polynomial function of the components of the Riemann tensor. If L also contains additive terms which are linear in or independent of these components, I shall call it an 'inhomogeneous quadratic Lagrangian'.

Every quadratic Lagrangian density $\mathfrak{L}$ is the sum, with constant coefficients, of scalar densities $\mathfrak{I}$ of the form

$$w g_{mn} g^{\cdot\cdot} g^{\cdot\cdot} g^{\cdot\cdot} R^{\cdot}{}_{jkl} R^{\cdot}{}_{bcd}, \qquad \epsilon^{\cdot\cdot\cdot\cdot} g_{mn} g^{\cdot\cdot} R^{\cdot}{}_{jkl} R^{\cdot}{}_{bcd},$$

where the set of superscripts, indicated by dots, is a permutation of the set of subscripts, $R^{i}{}_{jkl}$ depending on the components of the connection and their first derivatives alone. It is obvious by inspection that every $\mathfrak{I}$ is invariant under 'conformal transformations', i.e. unaffected by the substitution $g_{ij} \to \phi g_{ij}$, where ϕ is an arbitrary scalar function. Since g_{ij} and $\Gamma^{l}{}_{mn}$ are initially quite unrelated, it follows that $\mathfrak{A}^{ij}$ and $\mathfrak{B}_{k}{}^{ij}$ are likewise conformally invariant. Accordingly, if g_{ij} is any solution of the P-equations generated by any quadratic Lagrangian, then so is ϕg_{ij}. (By contrast, the solution of the g-equations do not admit this arbitrariness, apart from one special case.)

It may be noted that the conclusions just drawn hinge on the dimensional number being 4. Were it $2n$ (n integral), corresponding results would obtain for Lagrangians which are homogeneous functions (not necessarily polynomial functions) of degree n of the components of the Riemann tensor.

3. The Lagrangian R^2

The P-equations are in this case

$$R(R^{(ij)} - \tfrac{1}{4} g^{ij} R) = 0, \tag{3.1}$$

$$(\mathfrak{g}^{ij} R)_{;k} - \delta^{(j}{}_{k} (\mathfrak{g}^{i)l} R)_{;l} = 0. \tag{3.2}$$

Both equations are satisfied when $R = 0$. Thus only a single condition is imposed upon the 40 functions $\Gamma^{k}{}_{ij}$, while the g_{ij} remain entirely arbitrary. This trivial situation is henceforth excluded: $R \neq 0$.

By contraction (3.2) leads to $(\mathfrak{g}^{ij} R)_{;j} = 0$ and the equation reduces to

$$(\mathfrak{g}^{ij} R)_{;k} = 0.$$

Writing this out in full and lowering indices, there comes

$$g_{ij,k} - \Gamma^{l}{}_{ik} g_{lj} - \Gamma^{l}{}_{jk} g_{li} - g_{ij}[(\ln w)_{,k} + (\ln R)_{,k} - \Gamma_{k}] = 0, \tag{3.3}$$

where $\Gamma_k =: \Gamma^{l}{}_{lk}$. By transvection of (3.3) with g_{ij} it follows that

$$\Gamma_k = (\ln w)_{,k} + 2(\ln R)_{,k}.$$

Thus Γ_k is a gradient, which implies that R_{ij} is symmetric, since in any A_n $R_{[ij]} = \Gamma_{[i,j]}$. Furthermore (3.3) reduces to

$$\bar{g}_{ij,k} = \Gamma^{l}{}_{ik} \bar{g}_{lj} + \Gamma^{l}{}_{jk} \bar{g}_{li}, \tag{3.4}$$

where $\bar{g}_{ij} := R g_{ij}$. At once,

$$\Gamma^{k}{}_{ij} = \tfrac{1}{2} \bar{g}^{kl} (\bar{g}_{li,j} + \bar{g}_{lj,i} - \bar{g}_{ij,l}). \tag{3.5}$$

($\bar{g}^{kl}$ is the conjugate of $\bar{g}_{ij}$. Therefore $\bar{g}^{kl} = R^{-1} g^{kl}$.)

Equation (3.5) shows that the connection of the A_4^* is that of an auxiliary Riemann space $\bar{V}_4$, the metric tensor of which is $\bar{g}_{ij}$; that is to say, the connection is metric with respect to $\bar{g}_{ij}$. At the same time (3.1) now reads

$$R_{ij} = \tfrac{1}{4}\bar{g}_{ij}. \tag{3.6}$$

Bearing in mind that R_{ij} is a function only of the connection and its first derivative, R_{ij} is at the same time the Ricci tensor of the $\bar{V}_4$; and according to (3.6) the latter is an Einstein space of unit scalar curvature $\bar{R}$ ($:= \bar{g}^{ij}R_{ij}$).

Let the class of such Einstein spaces $\bar{V}_4$ be regarded as known, i.e. regard the $\bar{g}_{ij}$ as known functions. Then $R = g^{ij}R_{ij} = \frac{1}{4}g^{ij}\bar{g}_{ij}$, so that the equation

$$\bar{g}_{ij} = \tfrac{1}{4}g_{ij}g^{mn}\bar{g}_{mn} \tag{3.7}$$

determines g_{ij}. It shows that g_{ij} is conformal to $\bar{g}_{ij}$, consistently with the result of § 2. At any rate, granted that $R \neq 0$, the general solution of the P-equations (3.1) and (3.2) is any Riemann space conformal to an Einstein space.

4. The Lagrangian $R^{kl}R_{(kl)}$

Write $R_{(kl)} := P_{kl}$, $R_{[kl]} := F_{kl}$. Then to the Lagrangian $R_{ij}R^{ij}$ in V_4 there corresponds a whole family of Lagrangians in A_4^*, namely

$$L = P_{kl}P^{kl} + \alpha F_{kl}F^{kl}, \tag{4.1}$$

where α is a constant which can be chosen at will. The general case leads to various difficulties, and it will suffice here to concentrate on the relatively tractable case $\alpha = 0$. Then the P-equations are

$$\mathfrak{P}^{ij}{}_{;k} - \delta^{(i}{}_k\mathfrak{P}^{j)l}{}_{;l} = 0, \tag{4.2}$$

$$P^{im}P^j{}_m - \tfrac{1}{4}g^{ij}P^{mn}P_{mn} = 0. \tag{4.3}$$

By contracting (4.2) one infers that $\mathfrak{P}^{ij}{}_{;j}$ must vanish, so that this equation reduces to

$$\mathfrak{P}^{ij}{}_{;k} = 0, \tag{4.4}$$

or, written out in full,

$$P^{ij}{}_{,k} + \Gamma^i{}_{kl}P^{lj} + \Gamma^j{}_{kl}P^{li} + [(\ln w)_{,k} - \Gamma_k]P^{ij} = 0. \tag{4.5}$$

Now (4.2) and (4.3) are both satisfied when $P_{ij} = 0$, i.e. if only the A_4^* is Ricci-flat, the g_{ij} remaining entirely arbitrary. This somewhat trivial possibility is therefore set aside in what follows. However, I shall make the stronger assumption that P^{ij} is of full rank. (It is worth noting that in an A_n^* ($n \neq 4$) this assumption would be inconsistent with (4.3).) Then since, by assumption, g_{ij} is of full rank,

$$p := \det P_{ij} = w^4 \det P^{ij} \neq 0, \tag{4.6}$$

and there exists a tensor $\mathbf{P}_{ij}$ such that

$$\mathbf{P}_{ij}P^{kj} = \delta_i^k. \tag{4.7}$$

Transvection of (4.5) with $\mathbf{P}_{ij}$ shows that

$$\Gamma_k = (\tfrac{1}{2}\ln p)_{,k}, \tag{4.8}$$

since, bearing (4.6) in mind, $\mathbf{P}_{ij}P^{ij}{}_{,k} = [\ln(w^{-4}p)]_{,k}$. The fact that Γ_k is a gradient implies incidentally that $F_{ij}=0$. Now transvect (4.5) with $\mathbf{P}_{im}\mathbf{P}_{jn}$, and then

$$\Gamma^k{}_{il}\mathbf{P}_{jk}+\Gamma^k{}_{jl}\mathbf{P}_{il}=\mathbf{P}_{ij,k}+(\tfrac{1}{2}\ln p-\ln w)_{,k}\mathbf{P}_{ij}. \tag{4.9}$$

Let

$$\gamma_{ij}:=w^{-1}\sqrt{p}\mathbf{P}_{ij}, \qquad \gamma^{ij}:=P^{ij}w/\sqrt{p}, \tag{4.10}$$

so that $\gamma_{ik}\gamma^{jk}=\delta_i^j$. Then (4.9) reduces to

$$\Gamma^k{}_{il}\gamma_{jk}+\Gamma^k{}_{jl}\gamma_{ik}=\gamma_{ij,k},$$

whence

$$\Gamma^k{}_{ij}=\tfrac{1}{2}\gamma^{kl}(\gamma_{li,j}+\gamma_{lj,i}-\gamma_{ij,l}). \tag{4.11}$$

Accordingly the connection of the A_4^* is that of an auxiliary Riemann space $\tilde{V}_4$ whose metric tensor is γ_{ij}. Moreover, (4.3) shows that P_{ij} is a scalar multiple of $\mathbf{P}_{ij}$, i.e. of γ_{ij}, so that $\tilde{V}_4$ is an Einstein space. Equation (4.3) may be rewritten in the equivalent form

$$g^{kl}\gamma_{ik}\gamma_{jl}=\tfrac{1}{4}g_{ij}g^{kl}g^{mn}\gamma_{km}\gamma_{ln}. \tag{4.12}$$

As in § 3 the γ_{ij} are to be regarded as known functions, so that (4.12) is an equation for g_{ij}. It is satisfied if $g_{ij}=\phi\gamma_{ij}$, where ϕ is an arbitrary function, i.e. granted that $p\neq 0$ the P-equations are satisfied by any Riemann space conformal to an Einstein space. However, whereas in § 3 this was the only solution of the P-equation (when $R\neq 0$), this is not the case here. To see this it suffices to consider the case where γ_{ij} is diagonal. Then (4.12) is satisfied by

$$g_{ij}=\eta_i\phi\gamma_{ij} \qquad \text{(not summed)}, \tag{4.13}$$

where ϕ is arbitrary and, for each value of i, $\eta_i^2=1$. The possibility of freely choosing the signs of the η_i allows one to give g_{ij} the correct signature whatever that of γ_{ij} may be.

5. The Lagrangian $R_{[ij]}R^{ij}$

In § 4 only the Lagrangian formed from the symmetric part of R_{ij} was examined in detail. It is worth considering, as a counterpart, the curious case in which L contains only the skew-symmetric part. Thus, taking $L=F_{ij}F^{ij}$, one can virtually read the P-equations off from (4.2) and (4.3). They are in effect

$$\mathfrak{F}^{ij}{}_{;j}=0, \tag{5.1}$$

$$F^{im}F^j{}_m-\tfrac{1}{4}g^{ij}F^{mn}F_{mn}=0. \tag{5.2}$$

If one formally identifies F^{ij} with the field tensor of the Maxwell field, so that $-\frac{1}{2}\Gamma_i$ represents the electromagnetic potential, then (5.1) and (5.2) are the equations for a field whose energy–momentum tensor vanishes. This, however, implies that $F_{ij}=0$. The P-equations therefore merely require the A_4^* to be such that Γ_i is a gradient. They thus impose only four conditions upon the 40 functions $\Gamma^m{}_{kl}$, and they say nothing at all about the g_{ij}.

Confronted with such peculiarities one may of course take the view that one should simply exclude from the outset Lagrangians which give rise to them. After all, in the context of g-variations one excludes $L':=\delta_{abcd}{}^{klmn}R^{ab}{}_{kl}R^{cd}{}_{mn}$ as a Lagrangian since it gives rise to no g-equations at all, L' being, as is well known, functionally constant in a V_4 (e.g. Buchdahl 1970). In other words, the variational principle $\delta\int L' w \, dx = 0$ establishes no control over the g_{ij} at all.

6. Inhomogeneous quadratic Lagrangians

It will suffice to consider the special case

$$L = a + bR + R^2 \qquad (a, b = \text{constant}). \tag{6.1}$$

The P-equations are

$$(L'\mathfrak{g}^{ij})_{;k} - \delta^{(i}{}_k (L'\mathfrak{g}^{j)l})_{,l} = 0, \tag{6.2}$$

$$R^{(ij)}L' - \tfrac{1}{2}g^{ij}L' = 0, \tag{6.3}$$

where $L' := dL/dR$. By contraction (6.3) requires that

$$bR + 2a = 0. \tag{6.4}$$

The case $a = b = 0$ is excluded here since L is to be inhomogeneous. This leaves the four alternatives (i) $a \neq 0$, $b = 0$, (ii) $a = 0$, $b \neq 0$, (iii) $a \neq 0$, $b \neq 0$, $b^2 - 4a \neq 0$, (iv) $a \neq 0$, $b \neq 0$, $b^2 - 4a = 0$ to be considered.

Case (*i*). One has a contradiction, the condition $a \neq 0$ being in conflict with (6.4), i.e. $\mathfrak{L}$ cannot be extremised under P-variations at all.

Case (*ii*). According to (6.4) R must vanish and L in effect reduces to R, so that A_4^* is a Ricci-flat Riemann space.

Case (*iii*). This is also somewhat trivial in as far as R must be constant ($= -2a/b$), and one is led to a Riemannian Einstein space.

Case (*iv*). R must again be constant, but beyond this the P-equations become nugatory, i.e. the only condition imposed upon the A_4^* is that it have constant scalar curvature.

It is striking how different the various conclusions just drawn are from those to which one is led under g-variations, these giving rise to the equation

$$R_{:ij} + (R + \tfrac{1}{2}b)R_{ij} - \tfrac{1}{4}g_{ij}(4\Box R + R^2 + bR + a) = 0, \tag{6.5}$$

whence, by contraction,

$$6\Box R + bR + 2a = 0. \tag{6.6}$$

In no case do these equations require that R be necessarily constant, nor do they lead to a contradiction when $b = 0$.

Acknowledgment

I wish to express my thanks to Hubert Goenner for helpful discussions.

References

Buchdahl H A 1960 *Proc. Camb. Phil. Soc.* **56** 396–400
—— 1970 *Proc. Camb. Phil. Soc.* **68** 179–185