First-order formalism treatment of $R+R^2$ gravity

Bahman Shahid-Saless
Department of Physics, Campus Box 390, University of Colorado, Boulder, Colorado 80309
(Received 27 November 1985)

We show that the field equations of $R+R^2$ gravity formulated via the first-order formalism are different from those derived in the second-order formalism. These new field equations lead to a new set of connection coefficients that are conformally metric and in general not fully metric compatible. This conformally metric theory is unique in the sense that the conformal gauge field is readily identifiable in terms of the trace of the energy-momentum tensor and its derivatives. This identification leads to a new set of differential equations that can be interpreted as evolution equations for cosmological quantities. The importance of these equations in the early Universe is discussed.

INTRODUCTION

The idea of extending the Einstein-Hilbert action of gravity to include higher-derivative terms has been around for decades.[1–4] Recently the motivation for examining such extended Einstein's theories (EET's) has been due to their renormalizability,[5] and their being ghost-free. They have also been predicted as the low-energy limit of the heterotic superstring theory.[6]

It is well known that these EET's satisfy the "crucial" tests of general relativity (GR) in empty space. Nieh and Rauch[7] have found a class of them which are not renormalizable but satisfy unitarity and Birkhoff's theorem (BT). This has lead many to investigate their properties in regions where the energy-momentum tensor does not vanish. These investigations, however, have used a second-order formalism where the connections are considered to be the metric-compatible connections of GR—the Christoffel symbols (CS's). In this paper we assume arbitrary connections and vary the action as if the metric and the connections are independent fields, i.e., via the first-order, "Palatini" method. We arrive at a new set of field equations, different from those obtained via second-order techniques.

Simple analysis of the new connections reveals that this theory is a subclass of "conformally metric" theories in which the covariant derivative of the metric is proportional to the metric by a vector field. Weyl was the first to consider such theories.[8] Smolin[9,10] has considered these theories in the light of the possibility of the existence of conformal symmetry in the early Universe. This symmetry would then be broken at later stages in the evolution of the Universe. In this paper, the Weyl field can be identified in terms of the trace of the energy-momentum tensor and its derivatives. Furthermore, new conserved quantities can be derived from the new Bianchi identities. These are the modified energy-momentum tensor components.

It is worth mentioning that in standard GR, the Palatini formalism and the second-order formalism yield the same results. The former, though, because of its generality is even more appealing when one considers gravitation in the quantum era. In fact there is no *a priori* reason to assume any dependence between the metric and the connections if the strong principle of equivalence is not assumed to be true. We will consider the implications of the existence of a nonvanishing Weyl tensor on the strong equivalence principle.

In Sec. I of this paper, we start with a quick review of Weyl's conformal metric theory. A more detailed treatment can be found in Ref. 9. In Sec. II, we construct and vary our action to derive the field equations and in Sec. III, the connections. In Sec. IV we discuss the transformation properties of the connections and their implications on Einstein's equivalence. Section V will be devoted to our new differential equations and their interpretations. Cosmological solutions to these equations will be the subject of a future paper. A recipe for calculations will be the subject of Sec. VI and our conclusions, the subject of the last section. The metric signature is $(-1,1,1,1)$ and we use Weinberg's[11] conventions.

I. WEYL'S CONFORMAL GEOMETRY

In the Weyl conformal geometry one abandons the requirement that the covariant derivative of the metric tensor vanishes. Instead one considers that the covariant derivative of the metric is proportional to itself by a vector,

$$\nabla_\alpha g_{\mu\nu}=b_\alpha g_{\mu\nu} , \tag{1.1}$$

where b_α is the Weyl field. Under a conformal rescaling of the metric given by

$$\tilde{g}_{\mu\nu}=g_{\mu\nu}\Omega^2(x) \tag{1.2}$$

we get

$$\nabla_\alpha\tilde{g}_{\mu\nu}=\tilde{b}_\alpha\tilde{g}_{\mu\nu} . \tag{1.3}$$

The vector field, under such transformation, transforms like

$$\tilde{b}_\lambda=b_\lambda+2\Omega^{-1}(x)\partial_\lambda\Omega(x) . \tag{1.4}$$

Equation (1.1) can be solved for the connections in terms of $g_{\mu\nu}$ and b_α to give

$$\Gamma^{\alpha}{}_{\mu\nu}=\begin{Bmatrix}\alpha\\ \mu\nu\end{Bmatrix}-\tfrac{1}{2}(\delta^{\alpha}{}_{\mu}b_{\nu}+\delta^{\alpha}{}_{\nu}b_{\mu}-g_{\mu\nu}b^{\alpha}) , \quad (1.5)$$

where $\{^{\alpha}_{\mu\nu}\}$ are the usual CS's of GR.

One can then define a conformally invariant curvature tensor, the Ricci tensor and the curvature scalar by the usual definitions and one gets

$$R^{\lambda}{}_{\mu\nu\kappa}=\partial_{\kappa}\Gamma^{\lambda}{}_{\mu\nu}-\partial_{\nu}\Gamma^{\lambda}{}_{\mu\kappa}+\Gamma^{\eta}{}_{\mu\nu}\Gamma^{\lambda}{}_{\kappa\eta}-\Gamma^{\eta}{}_{\mu\kappa}\Gamma^{\lambda}{}_{\nu\eta} , \quad (1.6)$$

$$R_{\mu\nu}=R^{(0)}_{\mu\nu}-\tfrac{3}{2}D_{\mu}b_{\nu}+\tfrac{1}{2}D_{\nu}b_{\mu}-\tfrac{1}{2}g_{\mu\nu}D\cdot b$$
$$-\tfrac{1}{2}b_{\mu}b_{\nu}+\tfrac{1}{2}g_{\mu\nu}b^{2} , \quad (1.7)$$

$$R=R^{(0)}-3D\cdot b+\tfrac{3}{2}b^{2} . \quad (1.8)$$

Here D_{μ} is the covariant derivative associated with CS's, and a superscript zero in parentheses denotes the metric-compatible quantities:

$$D_{\alpha}g_{\mu\nu}\equiv 0 . \quad (1.9)$$

The Bianchi identity

$$\nabla_{\rho}R^{\lambda}{}_{\mu\nu\kappa}+\nabla_{\kappa}R^{\lambda}{}_{\mu\rho\nu}+\nabla_{\nu}R^{\lambda}{}_{\mu\kappa\rho}=0 \quad (1.10)$$

still holds; however, the identity

$$R^{(0)\alpha\beta}{}_{\mu\nu}=-R^{(0)\beta\alpha}{}_{\mu\nu} \quad (1.11)$$

does not hold in general. It is replaced by

$$R^{\alpha\beta}{}_{\mu\nu}+R^{\beta\alpha}{}_{\mu\nu}=H_{\mu\nu}g^{\alpha\beta} , \quad (1.12)$$

where $H_{\mu\nu}=\partial_{\mu}b_{\nu}-\partial_{\nu}b_{\mu}$.

Also one finds that

$$R_{\mu\nu}-R_{\nu\mu}=-2H_{\mu\nu} . \quad (1.13)$$

II. DERIVATION OF THE FIELD EQUATIONS

A. The action

We choose the simplest action that satisfies Birkhoff's theorem and has no torsion. [In general all combinations of $R^2(R^2, R_{\mu\nu}R^{\mu\nu}, R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta})$ can be included. Topological invariance eliminates one. The other two are dependent for isotropic and homogeneous models.] We also limit our analysis to matter fields for which the Lagrangian density does not involve the space-time connection explicitly. The inclusion of a cosmological constant at this time is avoided since addition of such a term in the field equations is straightforward. The action is given by

$$A=\int\sqrt{-g}\,(R+\alpha R^{2}+\mathcal{L}_{\text{matter}})d^{4}x . \quad (2.1)$$

B. The field equations

Varying the action with respect to the metric $g_{\mu\nu}$ gives

$$G_{\mu\nu}+2\alpha R\,(R_{\mu\nu}-\tfrac{1}{4}g_{\mu\nu}R)=KT_{\mu\nu} , \quad (2.2)$$

where $G_{\mu\nu}$ is the Einstein's tensor, $k=-8\pi G$, and α a parameter as yet unknown that has units of $1/(\text{mass})^2$. [In the quantum era, one expects α to be of the order of $1/(m_{\text{Planck}})^2$.]

Note that R is identified with its variational properties in a geodesic frame. That is,

$$\delta R_{\mu\nu}=\nabla_{\rho}(\delta\Gamma^{\rho}{}_{\mu\nu})-\nabla_{\nu}(\delta\Gamma^{\rho}{}_{\mu\rho}) . \quad (2.3)$$

Note that these equations differ from the ones derived via second-order formalism.[12–15] They are void of covariant derivatives of R and the trace of the field equations is identical with that of GR:

$$R=-KT . \quad (2.4)$$

III. THE CONNECTIONS

Varying the action with respect to connections gives

$$\nabla_{\mu}[\sqrt{-g}\,g^{\alpha\beta}(1+2\alpha R)]=0 . \quad (3.1)$$

After some manipulations, this equation yields

$$\nabla_{\alpha}g_{\mu\nu}=\frac{-2\alpha R_{,\alpha}}{1+2\alpha R}g_{\mu\nu} , \quad (3.2)$$

$$\nabla_{\alpha}g^{\mu\nu}=\frac{2\alpha R_{,\alpha}}{1+2\alpha R}g^{\mu\nu} . \quad (3.3)$$

Comparing the above equations with Eq. (1.1) helps identify this equality with that for conformal invariance. That is, one identities $-2\alpha R_{,\alpha}/(1+2\alpha R)$ with the Weyl field and the connections are simply

$$\Gamma^{\sigma}{}_{\mu\nu}=\begin{Bmatrix}\sigma\\ \mu\nu\end{Bmatrix}+\frac{\alpha}{1+2\alpha R}(\delta^{\sigma}{}_{\mu}R_{,\nu}+\delta^{\sigma}{}_{\nu}R_{,\mu}-g_{\mu\nu}g^{\sigma\alpha}R_{,\alpha}) . \quad (3.4)$$

The trace equation (2.4) tells us

$$b_{\mu}=\frac{-2\alpha R_{,\mu}}{1+2\alpha R}=\frac{2\alpha K}{1-2\alpha KT}T_{,\mu} . \quad (3.5)$$

So the Weyl field is nonzero wherever the trace of the energy-momentum tensor varies with respect to the coordinates.

IV. TRANSFORMATION PROPERTIES OF THE CONNECTIONS

In the derivation of the field equations, we assumed the existence of a geodesic frame, a frame where the connections vanish. Here we construct such a frame by noting that the connections transform just like the CS's. This is because the modifications to the CS's are tensors. Thus, under a coordinate transformation,

$$\Gamma'^{\sigma}{}_{\mu\nu}=\Gamma^{\alpha}{}_{\rho\delta}\left[\frac{\partial x^{\rho}}{\partial x'^{\mu}}\frac{\partial x^{\delta}}{\partial x'^{\nu}}\frac{\partial x'^{\sigma}}{\partial x^{\alpha}}\right]+\frac{\partial x'^{\sigma}}{\partial x^{\beta}}\frac{\partial^{2}x^{\beta}}{\partial x'^{\mu}\partial x'^{\nu}} . \quad (4.1)$$

A quadratic transformation given by

$$x^{\mu}=x'^{\mu}-x'^{\mu}|_{p}+\tfrac{1}{2}A^{\mu}{}_{\kappa\lambda}(x'^{\kappa}-x'^{\kappa}|_{p})(x'^{\lambda}-x'^{\lambda}|_{p}) \quad (4.2)$$

causes the connections to vanish, provided that

$$A^{\sigma}{}_{\mu\lambda}=\Gamma^{\sigma}{}_{\mu\lambda}|_{p} . \quad (4.3)$$

The vanishing of the connections in a particular frame however does not mean that the metric is Minkowskian there, because the covariant derivative of the metric is not in general zero, and, therefore, from Eqs. (3.2) and (3.3),

$$\partial_\alpha g_{\mu\nu}=b_\alpha g_{\mu\nu} \; . \tag{4.4}$$

Therefore the strong principle of equivalence is in general disobeyed. The frame in which the connections vanish is not necessarily the frame in which the metric is flat. If the parameter α is as small as one expects, nonequivalence cannot be measured at the present time for any known cosmological or astrophysical phenomenon. It could only be observed at the Planck time.

With the identification of b_μ, one can easily show that $H_{\mu\nu}$ vanishes, and one has

$$R_{\mu\nu}=R_{\nu\mu} \; .$$

This, however, may seem to trivialize the theory. $H_{\mu\nu}$ is gauge invariant and will always be zero. Furthermore, one can always gauge the Weyl field away by a conformal gauge transformation and the vanishing of both b_μ and $H_{\mu\nu}$ would lead us back to GR. It is therefore necessary to include matter which is not invariant under conformal transformations, at least classically. Equation (3.5) would then prevent the gauge field from vanishing by an arbitrary conformal gauge transformation.

In the quantum era, the problem becomes more interesting. Even if the whole theory, including the matter, is conformally invariant at the classical level, renormalization effects bring about anomalies which breakdown the existing conformal symmetry. This breakdown manifests itself in the form of the well-known anomalous trace of the energy-momentum tensor. After such a symmetry breakdown, the Weyl field becomes nontrivial and cannot be set to zero. Equation (4.4) would then imply that the existence of a conformal anomaly results in a violation of the strong equivalence principle. Since anomalies are quantum-mechanical effects, violations would be of the order of Planck's constant or smaller (there is a factor of α in front of $T_{\mu\nu}$).

V. DIFFERENTIAL EQUATIONS

From the field equations (2.2) and the trace equation (2.4) one can express the Ricci tensor in terms of the energy-momentum tensor:

$$R_{\mu\nu}=\frac{1}{1-2\alpha KT}(KT_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}KT+\tfrac{1}{2}\alpha g_{\mu\nu}K^2T^2) \tag{5.1}$$

and similarly the Einstein tensor,

$$G_{\mu\nu}=\frac{K}{1-2\alpha KT}(T_{\mu\nu}-\tfrac{1}{2}\alpha g_{\mu\nu}KT^2) \; . \tag{5.2}$$

Just as in GR these quantities are defined once the distribution of matter-energy is given. The Bianchi identity, (1.10), however, with the definition of the covariant derivative of the metric, (1.1), can serve to identify the conserved quantities. Starting with (1.10), one can show that

$$\nabla_\alpha G^\alpha{}_\mu=-b_\alpha G^\alpha{}_\mu \tag{5.3}$$

and, therefore,

$$\nabla_\mu\left[\frac{G^\mu{}_\nu}{1-2\alpha KT}\right]\equiv 0 \; . \tag{5.4}$$

We define a new energy-momentum tensor that is now conserved:

$$\mathscr{T}_{\mu\nu}=\frac{1}{(1-2\alpha KT)^2}(T_{\mu\nu}-\tfrac{1}{2}\alpha g_{\mu\nu}KT^2) \; . \tag{5.5}$$

Looking at Eqs. (1.8) and (2.4) we can form a differential equation,

$$-KT=R^{(0)}-3D\cdot b(T)+\tfrac{3}{2}b^2(T) \; , \tag{5.6}$$

where b is now a function of T and its derivatives as defined by Eq. (3.5). Similarly Eqs. (1.7) and (5.1) give

$$R^{(0)}_{\mu\nu}-D_\mu b_\nu-\tfrac{1}{2}g_{\mu\nu}D\cdot b-\tfrac{1}{2}b_\mu b_\nu+\tfrac{1}{2}g_{\mu\nu}b^2 =\frac{1}{1-2\alpha KT}(KT_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}KT+\tfrac{1}{2}\alpha g_{\mu\nu}K^2T^2) \; . \tag{5.7}$$

These differential equations must be satisfied at all times. They will serve as evolution equations for T and $T_{\mu\nu}$.

One must worry about the consistency of these equations. It can be shown that Eqs. (1.7) and (1.8) are in fact consistent with the Bianchi identity (5.3).

The above differential equations will tell how various cosmological quantities such as energy density, pressure, and temperature change as the Universe evolves from Planck time to the inflationary era. The solutions to these equations are presently under consideration. They will constitute the subject of a future paper.

VI. RECIPE FOR CALCULATIONS

Physical calculations are a bit more subtle in this theory. The metric itself does not supply all the tools one needs to study different cosmological models. The connections, the curvature tensor, Ricci tensor, and the curvature scalar are all dependent not only on the metric but on the form of the energy-momentum tensor and its coordinate derivatives. The matter-energy distribution therefore plays a more fundamental role in this theory than in standard GR.

Recently there have been some efforts to guess the form of the energy-momentum tensor around the Planck time.[16] It is believed that a straightforward generalization of our theory to ten dimensions and the use of the results of such works may help us find solutions of the evolution equations and lead us to interesting predictions about compactification and early Universe evolution. This is currently being pursued by the author.

It is also interesting to note that in the radiation-dominated era, the trace of the energy-momentum tensor vanishes if one takes the equation of state to be $p=\frac{1}{3}\rho$. Standard GR cosmology will hold at this time.

CONCLUSIONS

We have applied the Palatini formalism to an extended Einstein-Hilbert action (2.1). We have found the following.

(a) The field equations are different from those derived by using a second-order technique.

(b) The new theory is a subclass of conformally metric theories first introduced by Weyl.

(c) This theory has a unique geometry in which the Weyl field is a function of the curvature scalar—or the trace of the energy-momentum tensor—and its derivatives.

(d) Einstein's principle of strong equivalence is violated to leading order of α if the derivative of the trace of the energy-momentum tensor is nonzero. This is important especially when conformal symmetry breaks down. The existence of conformal anomaly results in the violation of the equivalence principle.

(e) A set of differential equations are found that describe the space-time evolution of the relevant cosmological quantities. These must be satisfied for the theory to be self-consistent.

Full implications of this model are yet unknown. Modifications to standard GR would manifest themselves most, where/when thermodynamic configurations are rapidly changing. Our only hopes to observe such effects would be either in astrophysical phenomena or in the early Universe.

ACKNOWLEDGMENTS

I thank my advisor Professor Neil Ashby for his continual support throughout this work and Dr. F. Graziani for helpful discussions about the early Universe. Also I would like to thank Philip Ensign for many helpful suggestions.

[1]H. Weyl, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. **Kl**, 465 (1918); Ann. Phys. (Leipzig) **59**, 101 (1919); Phys. Z. **22**, 473 (1921).

[2]W. Pauli, Phys. Z. **20**, 457 (1919).

[3]A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge University, Cambridge, England, 1965), p. 141.

[4]C. N. Yang, Phys. Rev. Lett. **33**, 445 (1974).

[5]K. S. Stelle, Phys. Rev. D **16**, 953 (1977).

[6]See, for example, B. Zwiebach, Phys. Lett. **156B**, 315 (1985).

[7]H. T. Nieh and R. Rauch, Phys. Lett. **81A**, 113 (1981).

[8]H. Weyl, *Space-Time Matter* (Dover, New York, 1922), Sec. 35; Sitzungsber. Preuss. Akad. Wiss. Phys. Math. **K1**, 465 (1918), reprinted in *The Principles of Relativity* (Dover, New York, 1923).

[9]L. Smolin, Nucl. Phys. **B160**, 253 (1979).

[10]L. Smolin, Nucl. Phys. **B247**, 511 (1984).

[11]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).

[12]T. V. Ruzmaikina and A. A. Ruzmaikin, Zh. Eksp. Teor. Fiz. **57**, 680 (1969) [Sov. Phys. JETP **30**, 372 (1970)].

[13]K. Tomita, T. Azuma, and H. Nariai, Prog. Theor. Phys. **60**, 403 (1978).

[14]H. Nariai, Prog. Theor. Phys. **46**, 433 (1971).

[15]H. Nariai, Report No. RRK 85-20, 1985 (unpublished).

[16]M. Gleiser and J. G. Taylor, Phys. Lett. **164B**, 36 (1985).