Yang-Mills formulation of gravitational dynamics

Edward E. Fairchild, Jr.*

Department of Astrophysics, Oxford University, Oxford, England OX1 3RQ and McDonnell Center for the Space Sciences, Washington University, St. Louis, Missouri 63130†

(Received 4 October 1976; revised manuscript received 29 March 1977)

A Yang-Mills formulation of the dynamics of the gravitational field is given based on gauged spin transformations acting on a Dirac spin space. The gravitational field is carried by the spin connection following the point of view of Schrödinger, Laurent, and Loos. Differing from them, the dynamical statement is in a strictly Yang-Mills form. The Lagrangian for the model theory has a single massive Dirac field as source and is one investigated by Leutwyler but with a different end in mind. I compare it with Einstein's theory and show acceptable observational behavior for the theory in the usual circumstances. The theory is found to make contact with theories with torsion. A comparison between this work and previous work by Loos and Treat, Yang, Camenzind, and Carmeli is made.

I. INTRODUCTION

With the many theoretical successes of Yang-Mills–type gauge theories, the hope for a true unified field theory of fundamental interactions has been rekindled. The spark of this rekindling is noting that electromagnetism, gravity, and the new theories of weak interactions are all of the local gauge type.[1–4] In this paper, I will formulate the theory of gravity in the Yang-Mills form.[2]

I would like to take a moment to say what I mean by a Yang-Mills theory of gravity. There exists quite a bit of confusion in the literature in applying the term *gauge* to theories. I do not want to add to this confusion. I mean by a *gauge theory* any theory which has kinematics based on a local gauge group, gauge potential or linear connection, and gauge field or curvature. It is in this sense that both Maxwell's and Einstein's theories are gauge theories, for the gauge groups U(1) and $SO_0(3,1)$, respectively. I call a *Yang-Mills theory* one which is modeled on gauge kinematics but with dynamics structured after electromagnetism as Yang and Mills did for the gauge group SU(2). Thus, Einstein's theory, while a gauge theory is *not* a Yang-Mills theory; also, every Yang-Mills theory is a gauge theory but thus every gauge theory is not in Yang-Mills form. The point of this paper is to bring gravity into the Yang-Mills form.

This is not the first attempt at such a reformulation. Yang himself proposed a set of free gravitational equations based on this idea.[3] Also, Loos and Treat,[5] Camenzind,[6] and Carmeli[7] have proposed Yang-Mills formulations of gravity. The form I propose differs essentially from those of Camenzind and Carmeli, but can be thought of as a development of the ideas of Loos and Treat.

The important gauge group of any gauge theory is its gauge holonomy group.[8] For electromagnetism, it is U(1); for the original Yang-Mills theory, it is SU(2); for the Weinberg-Salam style theory of weak and electromagnetic interactions,[4] it is $SU(2)\otimes U(1)$; for Einstein's theory of gravity, it is $SO_0(3,1)$.[9] All essential kinematical dissimilarities between different gauge theories are known when one knows their different gauge holonomy groups and on which spaces the groups act. The central idea of any Yang-Mills reformulation is to keep the $SO_0(3,1)$ gauge kinematics of gravity while simultaneously discarding Einstein's equations as the primary dynamical statement. One then replaces them with some Yang-Mills style field equations. For the formulation discussed herein, though, I will make one change in kinematics. I will go over from the Lorentz group $SO_0(3,1)$ to its spin covering group $SL(2,C)$. In the Yang-Mills SU(2) theory, the gauge field is modeled (i.e., the gauge group acts) on isospinor fields; in Einstein's theory, the gauge field is modeled on tensor fields of the tangent space to space-time. My change $SO_0(3,1)\rightarrow SL(2,C)$ is the equivalent of going from a modeling on tensor fields to a modeling on spinor fields. Gravitation will be carried in the $SL(2,C)$ gauge field.

The history of gauged $SL(2,C)$ spin transformations predates by quite a bit the introduction of gauge fields into particle physics by Yang and Mills.[2] Schrödinger made the first definite contribution noting the possibility of carrying the gravitational and electromagnetic kinematics in the gauge geometry of the Dirac spin space.[10] Laurent transferred the usual Einstein dynamics into this framework.[11] Loos showed to what extent the spin space connection determines the usual Riemannian geometry of gravity uniquely.[12] Treat first noted a generality in the spin geometry overlooked by previous authors, a generality which plays an important part of this paper.[13] Leutwyler took a different point of view from all of these.[14] He attempted to identify the various parts of a general spin connection with meson fields for use in particle physics.[15] My point of view will be to keep the Schrödinger-Laurent-Loos idea of gravity in the spin geometry while looking at the dynamics from a Yang-Mills approach closely following Leutwyler.

I begin in the next section with a review both of Dirac fields on a curved space-time and of the gauged spin connection geometry. I will use this structure in a simple model of gravitation with a single massive Dirac source field in Sec. III. Section IV is concerned with reduction of the field equations to a form which looks more like those of Einstein's theory so that we can make a comparison. Observational viability of these equations is briefly examined. When we look at the nongravitational part of the spin geometry, we will find that we can make contact with theories with torsion. Finally, in the last section, I will summarize and relate this approach to other work on Yang-Mills gravity which I previously mentioned.

II. THE KINEMATICS OF GAUGED SPIN FIELDS

I intend to model this theory of gravity using the linear connection (i.e., gauge potential) associated with gauged spin transformations $S_a{}^b(x)$ on a Dirac spinor field ψ_a:

$$\psi_{\underline{a}} = S_{\underline{a}}{}^b(x)\psi_b \ . \tag{2.1}$$

The lower-case Latin indices $\{a, b, c, \ldots, g, h\}$ from the early part of the alphabet denote quantities which lie in the Dirac spin space over space-time. These indices take on the values 1–4. The bar under an index tells us it is in a different frame. I reserve the usual prime for other purposes. The Dirac spinors are identified with quantities having the index in the lower covariant position (ψ_a). The quantities with the index in the upper contravariant position ($\overline{\psi}^a$) are identified as the Dirac adjoint spinors since they transform with the inverse transformation $S^{-1}{}_a{}^{\underline{b}}(x)$.

The process of taking the Dirac adjoint $\psi_a \to \overline{\psi}^a$ defines a Hermitian metric tensor $g^{a\dot{b}}$ (with inverse $g_{a\dot{b}}$). The dot denotes transformation under the complex-conjugate spin transformation.[16] To form the Dirac adjoint of a spinor ψ_a, one complex-conjugates it to $\psi_{\dot{a}}$ and raises the index using $g^{a\dot{b}}$:

$$\overline{\psi}^a = g^{a\dot{b}}\psi_{\dot{b}} \ . \tag{2.2}$$

What makes spinors what they are is their 2–1 relationship with vectors, and consequently the 2–1 relationship between the groups $SO_0(3,1) \to SL(2,C)$. These two mappings are carried by the Dirac matrices which are quantities with both spinor and vector indices: $\gamma_{\mu a}{}^b$. I will use lower-case Greek indices from the middle of the alphabet (specifically $\{\kappa, \lambda, \mu, \nu, \pi, \rho, \sigma, \tau\}$) to denote event (i.e., tangent space) indices. They will take on the values 0–3. Given a spinor, one can now form a vector (its "current") via

$$v_\mu = \overline{\psi}^a \gamma_{\mu a}{}^b \psi_b \ . \tag{2.3}$$

Please note that in all of the following, I will be using a coordinate frame in the tangent space.[17] Also, I will often suppress spinor indices in a matrix convention. In this case, Eq. (2.3) would read

$$v_\mu \equiv \overline{\psi}\gamma_\mu\psi \ . \tag{2.4}$$

The important, defining property of the Dirac matrices is that they satisfy the Dirac relation

$$\gamma_{(\mu}\gamma_{\lambda)} = g_{\mu\lambda} I \ . \tag{2.5}$$

The $g_{\mu\lambda}$ is the event metric (i.e., metric of the tangent space of space-time) with time favoring signature -2. The I is the identity spin matrix with components $\delta_a{}^b$, the spin Kronecker δ. Since $g_{\mu\lambda}$ is a function of position in a general space-time, so also is the γ_μ.

Next we need linear connections on both the tangent space and the Dirac spin space. In the tangent space, we introduce the usual Christoffel connection $\Gamma_{\mu\kappa}{}^\lambda$ given in the usual way in terms of the metric.[18] In the spin space, we introduce a connection $\Phi_{\mu a}{}^b$. It defines a gauged spin-covariant derivative

$$\nabla_\mu \psi_a \equiv \partial_\mu \psi_a - \Phi_{\mu a}{}^b \psi_b \ , \tag{2.6}$$

which transforms homogeneously under gauged spin transformations. This implies the usual inhomogeneous transformation law of the connection

$$\Phi_\mu \to S\Phi_\mu S^{-1} - S\partial_\mu S^{-1} \ . \tag{2.7}$$

The action on adjoint spinors is

$$\nabla_\mu \overline{\psi}^a \equiv \partial_\mu \overline{\psi}^a + \Phi_{\mu b}{}^a \overline{\psi}^b \ .$$

This covariant derivative ∇_μ extends to mixed quantities in the obvious way using $\Gamma_{\mu\kappa}{}^\lambda$ on event indices. For example,

$$\nabla_\mu \gamma_\lambda = \partial_\mu \gamma_\lambda - \Gamma_{\mu\lambda}{}^\rho \gamma_\rho - [\Phi_\mu, \gamma_\lambda] \ . \tag{2.8}$$

It is common with internal gauge symmetries to expand the connection onto a complete basis of generators of the gauge group. The expansion coefficients are identified as the gauge fields. As with Einstein's theory, it is not convenient or illuminating to do that here

The noncommutativity of the covariant derivative leads, as usual, to curvatures or gauge fields for both $\Phi_{\mu a}{}^b$ and $\Gamma_{\mu\kappa}{}^\lambda$. The corresponding Ricci identities are

$$\nabla_{[\mu}\nabla_{\kappa]}\psi_a = -\tfrac{1}{2}\phi_{\mu\kappa a}{}^b \psi_b \tag{2.9a}$$

and

$$\nabla_{[\mu}\nabla_{\kappa]}v_\lambda = -\tfrac{1}{2}R_{\mu\kappa\lambda}{}^\rho v_\rho \ . \tag{2.9b}$$

For contravariant quantities $\bar{\psi}^a$ and w^λ,

$$\nabla_{[\mu}\nabla_{\kappa]}\bar{\psi}^a = \tfrac{1}{2}\phi_{\mu\kappa b}{}^a\bar{\psi}^b \tag{2.9c}$$

and

$$\nabla_{[\mu}\nabla_{\kappa]}w^\lambda = \tfrac{1}{2}R_{\mu\kappa\rho}{}^\lambda w^\rho \ . \tag{2.9d}$$

The curvatures are given in terms of the connections via

$$\phi_{\mu\kappa} \equiv 2\partial_{[\mu}\Phi_{\kappa]} - [\Phi_\mu, \Phi_\kappa] \tag{2.10a}$$

and

$$R_{\mu\kappa\lambda}{}^\rho \equiv 2\partial_{[\mu}\Gamma_{\kappa]\lambda}{}^\rho - \Gamma_{\mu\lambda}{}^\sigma\Gamma_{\kappa\sigma}{}^\rho + \Gamma_{\kappa\lambda}{}^\sigma\Gamma_{\mu\sigma}{}^\rho \ . \tag{2.10b}$$

They satisfy the Bianchi identities

$$\nabla_{[\mu}\phi_{\kappa\lambda]} = 0 \tag{2.11a}$$

and

$$\nabla_{[\mu}R_{\kappa\lambda]\rho}{}^\sigma = 0 \ . \tag{2.11b}$$

Just as the Dirac matrices, satisfying (2.5), link spinors to vectors and $SO_0(3,1)$ to $SL(2,C)$, the Dirac relation (2.5) links the two gauge connections. The form of the Riemann-Christoffel connection $\Gamma_{\mu\kappa}{}^\lambda$ (Ref. 18) implies and is implied by the metric condition

$$\nabla_\mu g_{\kappa\lambda} = 0 \ . \tag{2.12}$$

This constrains, via the Dirac relation (2.5), what the covariant derivative of γ_μ can be. One can show that (2.12) is satisfied if and only if

$$\nabla_\mu\gamma_\kappa = [M_\mu, \gamma_\kappa] \tag{2.13}$$

for arbitrary $M_{\mu a}{}^b$. This was first noted by Treat.[13] The proof of this as well as all further results of this section has been deferred to the Appendix.

If we consider the integrability condition of (2.13), we get a relation between the curvatures $\phi_{\mu\lambda a}{}^b$ and $R_{\mu\kappa\lambda}{}^\rho$,

$$\phi_{\mu\kappa} = \tfrac{1}{4}R_{\mu\kappa}{}^{\rho\sigma}\gamma_{\rho\sigma} - m_{\mu\kappa} \ , \tag{2.14}$$

where

$$m_{\mu\kappa} \equiv 2\nabla_{[\mu}M_{\kappa]} - [M_\mu, M_\kappa] \ . \tag{2.15}$$

The $\gamma_{\rho\sigma}$ are the double Dirac γ matrices defined by (A1) and are the complete set of anti-self-adjoint generators of $SL(2,C)$.

Finally, one may solve (2.13) directly to give Φ_μ in terms of $g_{\mu\lambda}$, M_μ, and γ_κ:

$$\begin{aligned}\Phi_\mu = {} & \tfrac{1}{4}\Gamma_{\mu\kappa}{}^\rho\gamma^\kappa{}_\rho - \tfrac{1}{16}\mathrm{tr}(\gamma^\rho\partial_\mu\gamma_\kappa)\gamma^\kappa{}_\rho \\ & - \tfrac{1}{96}\mathrm{tr}(\gamma^{\rho\sigma}\partial_\mu\gamma_\sigma)\gamma_\rho + \tfrac{1}{8}\mathrm{tr}(\gamma\partial_\mu\gamma_\kappa)\gamma^\kappa\gamma \\ & - \tfrac{1}{32}\mathrm{tr}(\gamma^\rho\gamma\partial_\mu\gamma_\rho)\gamma - M_\mu \ .\end{aligned} \tag{2.16}$$

The γ_μ, $\gamma_{\kappa\mu}$, $\gamma_\mu\gamma$, and γ are the complete set of 15 Dirac matrices defined in the Appendix.

III. THE DYNAMICS OF A MODEL GAUGED SPIN THEORY

Here I would like to pose a model theory using the gauged spin connection as the primary dynamical quantity. Let

$$\mathcal{L}_\Phi \equiv -\tfrac{1}{4}\omega^{-2}\mathrm{tr}(\phi_{\mu\kappa}\phi^{\mu\kappa}) + \tfrac{1}{2}\mu^2\omega^{-2}\mathrm{tr}(\phi_{\mu\kappa}\gamma^{\mu\kappa}) \tag{3.1}$$

be the free Lagrangian for the gauged spin connection Φ_μ. The Lagrangian is chosen to be as close to the Yang-Mills form as possible. It differs in the second term, which is lower order in the curvature $\phi_{\mu\kappa}$. One cannot construct such a first-order term for the general gauge case. This is one of a class of Lagrangians considered by Leutwyler for the spin connection.[14] The first term is just the usual Yang-Mills Lagrangian with dimensionless coupling constant ω. The second term is needed to give an acceptable Newtonian limit for the resulting gravity theory. The coupling constant μ has the dimension of mass.[19] The term truly acts much like a mass term since part of it is bilinear in Φ_μ as one can see using (2.10) and (A6).

As a source for the gauge field, I have chosen a single massive Dirac field minimally coupled to the Φ_μ field through the covariant derivative. Its Lagrangian is

$$\mathcal{L}_\psi \equiv \bar{\psi}(i\overleftrightarrow{\nabla} - mI)\psi \ . \tag{3.2}$$

I have used the notation

$$A\overleftrightarrow{\nabla}B \equiv \tfrac{1}{2}(A\gamma^\mu\nabla_\mu B - (\nabla_\mu A)\gamma^\mu B) \ .$$

Put these fields on some space-time manifold with metric $g_{\mu\kappa}$ and associated Riemann-Christoffel connection $\Gamma_{\mu\kappa}{}^\rho$.[18] The action is given by

$$\mathcal{A} \equiv \int d\mu\, g^{1/2}\mathcal{L} \tag{3.3}$$

with $\mathcal{L} \equiv \mathcal{L}_\Phi + \mathcal{L}_\psi$, $g \equiv -\det g_{\mu\lambda}$, and integration over space-time.

Varying Φ_μ, ψ, and $\bar{\psi}$ independently gives the field equations

$$\delta\Phi_\mu:\ \nabla_\kappa g^{1/2}(\phi^{\kappa\mu} - \mu^2\gamma^{\kappa\mu}) = -i\omega^2 g^{1/2}(\psi\bar{\psi}\gamma^\mu + \gamma^\mu\psi\bar{\psi}) \equiv -\omega^2 g^{1/2}j^\mu \ , \tag{3.4}$$

$$\delta\bar{\psi}:\ \tfrac{1}{2}i(\gamma^\mu\nabla_\mu + \nabla_\mu\gamma^\mu)\psi - m\psi = 0 \ , \tag{3.5}$$

and its complex conjugate. The first is the Yang-Mills-type field equation; the second is the Dirac equation coupled to Φ_μ. Expanding the current on the generators of the $SL(2,C)$ gauge group

$$j^\mu \equiv J^{\mu\kappa\lambda}\gamma_{\kappa\lambda} \tag{3.6}$$

shows us that

$$J^{\mu\kappa\lambda} = -\tfrac{1}{8} i g^{-1/2} \epsilon^{\mu\kappa\lambda\rho} \bar{\psi}\gamma_\rho \gamma \psi \ , \tag{3.7}$$

the spin density of the Dirac field. Other terms in the Clifford expansion of j^μ are taken to be zero since they are not coupled to any field.

Stopping at this point, we would have a theory for the spin connection in a fixed nondynamical metric background. To include dynamical gravity into a theory on a nondynamical background, one usually adds the Hilbert Lagrangian

$$\mathcal{L}_g = (16\pi G)^{-1} R, \quad R \equiv R_{\rho\kappa\mu}{}^{\rho} g^{\kappa\mu} \tag{3.8}$$

to the Lagrangian $\mathcal{L}$ and varies the metric as a dynamical quantity. *However,* I do not want to do this. The dynamics of gravity is already in the gauged spin Lagrangian $\mathcal{L}$. We have no need of adding any term such as (3.8), which is in fact what Leutwyler did in considering the spin connection.[14] Let us vary the metric in the action (3.3) *as it stands.* The field equation is

$$\delta g^{\mu\lambda}: \ g^{1/2}(W_{\mu\lambda} + T_{\mu\lambda}) = 0 \ , \tag{3.9}$$

where

$$W_{\mu\kappa} = -\omega^{-2} \mathrm{tr}(\phi_{(\mu}{}^{\lambda}\phi_{\kappa)\lambda} - \tfrac{1}{4} g_{\mu\kappa}\phi_{\rho\sigma}\phi^{\rho\sigma}) + \mu^2\omega^{-2}\mathrm{tr}(\phi_{(\mu}{}^{\lambda}\gamma_{\kappa)\lambda} - \tfrac{1}{2} g_{\mu\kappa}\phi_{\rho\sigma}\gamma^{\rho\sigma}) \tag{3.10}$$

and

$$2T_{\mu\kappa} = i\bar{\psi}\gamma_{(\mu}\nabla_{\kappa)}\psi - i(\nabla_{(\mu}\bar{\psi})\gamma_{\kappa)}\psi - g_{\mu\kappa}\bar{\psi}(i\bar{\nabla} - mI)\psi \tag{3.11}$$

are the usual symmetric stress-energy tensors derived from $\mathcal{L}_\Phi$ and $\mathcal{L}_\psi$, respectively:

$$W_{\mu\kappa} \equiv 2g^{-1/2}\frac{\delta \mathcal{Q}_\Phi}{\delta g^{\mu\kappa}} = 2g^{-1/2}\frac{\partial \mathcal{L}_\Phi}{\partial g^{\mu\kappa}} \ , \tag{3.12a}$$

$$T_{\mu\kappa} \equiv 2g^{-1/2}\frac{\delta \mathcal{Q}_\Phi}{\delta g^{\mu\kappa}} = 2g^{-1/2}\frac{\partial \mathcal{L}_\Phi}{\partial g^{\mu\kappa}} \ . \tag{3.12b}$$

The relations

$$\delta g^{1/2} = -\tfrac{1}{2} g_{\mu\lambda}\delta g^{\mu\lambda} \tag{3.13}$$

and

$$\delta\gamma^\kappa = \tfrac{1}{2}\gamma_\lambda \delta g^{\lambda\kappa} \tag{3.14}$$

are needed to carry out these variations. The feasibility of taking γ_μ as a function of $g_{\mu\lambda}$, as indicated in (3.14), is discussed in the Appendix.

I claim that the gravitational dynamics is contained in the tensor $W_{\mu\lambda}$. In the next section, I shall examine this claim.

IV. COMPARISON TO EINSTEIN'S THEORY

To compare the theory to Einstein's, one must transform (3.4) and (3.9) into a more familiar form. Substitute for $\phi_{\mu\kappa}$ from (2.14) and use (2.13) to evaluate the action of ∇_μ on the γ matrices. Field equation (3.4) becomes

$$\nabla_\kappa m^{\kappa\mu} + \mu^2[M_\kappa, \gamma^{\kappa\mu}] - \tfrac{1}{4}(\nabla_\kappa R^{\kappa\mu\rho\sigma})\gamma_{\rho\sigma} - \tfrac{1}{4} R^{\mu\kappa\rho\sigma}[M_\kappa, \gamma_{\rho\sigma}] = \omega^2 g^{1/2} j^\mu \ . \tag{4.1}$$

From (3.10), we can also rewrite $W_{\mu\kappa}$ as

$$\begin{aligned} W_{\kappa\lambda} = {} & \tfrac{1}{2}\omega^{-2}\mu^2(R_{\kappa\lambda} - \tfrac{1}{2} g_{\kappa\lambda} R) \\ & + \tfrac{1}{2}\omega^{-2}(R_\kappa{}^{\nu\rho\sigma}R_{\lambda\nu\rho\sigma} - \tfrac{1}{4} g_{\kappa\lambda} R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}) \\ & - \omega^{-2}(m_\kappa{}^\mu m_{\lambda\mu} - \tfrac{1}{4} g_{\kappa\lambda} m_{\rho\sigma} m^{\rho\sigma}) \\ & - \omega^{-2}\mu^2 \mathrm{tr}(m_{(\kappa}{}^{\mu}\gamma_{\lambda)\mu} - \tfrac{1}{2} g_{\kappa\lambda} m_{\rho\sigma}\gamma^{\rho\sigma}) \\ & + \tfrac{1}{2}\omega^{-2} R_{(\kappa}{}^{\mu\rho\sigma}\mathrm{tr}(m_{\lambda)\mu}\gamma_{\rho\sigma}) \\ & - \tfrac{1}{8} g_{\kappa\lambda} R^{\mu\nu\rho\sigma}\mathrm{tr}(m_{\mu\nu}\gamma_{\rho\sigma}) \\ \equiv {} & + \tfrac{1}{2}\omega^{-2}\mu^2 G_{\kappa\lambda} + \tfrac{1}{2}\omega^{-2} H_{\kappa\lambda} + V_{\kappa\lambda} \ . \end{aligned} \tag{4.2}$$

The first term is just the Einstein tensor $G_{\mu\lambda} \equiv R_{\mu\lambda} - \frac{1}{2} g_{\mu\lambda} R$. The second term $H_{\kappa\lambda}$ arises in the Yang gravitational theory.[3,20] The last three I have put into a stress-energy tensor $V_{\kappa\lambda}$ for the M_μ field. Now we can rewrite (3.9) as

$$G_{\mu\lambda} + 8\pi G\omega^{-2} H_{\mu\lambda} = -8\pi G(T_{\mu\lambda} + V_{\mu\lambda}) \ , \tag{4.3}$$

where one identifies $\frac{1}{2}\omega^{-2}\mu^2 = (8\pi G)^{-1}$.

There are two differences with Einstein's equations in (4.3). One is in the additional curvature term $H_{\mu\lambda}$, which is a modification to the free gravitational equations. The second is the source field M_μ which has not yet been identified. The field equation (4.1) is the field equation of this M_μ field. $V_{\mu\lambda}$ is its stress-energy tensor.

The vacuum gravitational equations are obtained by setting the fields ψ, $\bar{\psi}$, and M_μ to zero giving

$$G_{\mu\lambda} + 8\pi G\omega^{-2} H_{\mu\lambda} = 0 \tag{4.4}$$

and

$$\nabla_\kappa R^{\kappa\lambda\mu\rho} = 0 \tag{4.5}$$

for (4.3) and (4.1). One may verify from its definition that $H_{\mu\lambda}$ is trace-free. This and (4.4) imply that $R = 0$: Vacuum solutions are scalar-flat as with Einstein's theory. Expand the curvature in terms of the Weyl, Ricci, and scalar curvature [8]:

$$R_{\nu\kappa\lambda\mu} = C_{\nu\kappa\lambda\mu} - 2g_{[\nu[\lambda}R_{\kappa]\mu]} + \tfrac{1}{3} g_{[\nu[\lambda}g_{\kappa]\mu]} R \ . \tag{4.6}$$

Using the $R = 0$ condition, one finds that substituting into the definition of $H_{\mu\lambda}$ from (4.6) gives Eq. (4.3) as

$$R_{\kappa\lambda} - 16\pi G\omega^{-2} C_{\kappa\rho\lambda\sigma} R^{\rho\sigma} = 0 \ . \tag{4.7}$$

Contracting the Bianchi identity (2.11b) once and using (4.5) gives the equation

$$\nabla_{[\mu} R_{\kappa]\lambda} = 0 \ . \tag{4.8}$$

From these two equations, one can see that the Ricci-flat geometries $R_{\mu\lambda} = 0$ are vacuum solutions of these equations, just as in Einstein's theory. But are they the only solutions? The answer is yes. Equations (4.7) and (4.8) are equivalent to $R_{\mu\lambda}=0$. The proof of this, although direct, is a bit long and tedious and will be published elsewhere.[21]

To examine the Newtonian limit, expand the metric infinitesimally off of the flat Minkowski metric background $\eta_{\mu\lambda}$ to first order

$$g_{\mu\lambda}=\eta_{\mu\lambda}+h_{\mu\lambda}\ .$$

Since $H_{\mu\lambda}$ is quadratic in the curvature, it is only of second order in $h_{\mu\lambda}$ and can be ignored. Identifying the total mass density $\rho\equiv T_{00}+V_{00}$ in (4.3) gives the Newtonian limit following the usual proof used in Einstein's theory.[22]

The only loose end remaining is to give M_μ an interpretation. It has a simple one in terms of the tangent-space kinematics. The spin gauge holonomy group is SL(2, C). Since the spin curvature $\phi_{\mu\kappa}$ can always be expanded in the generators of the gauge group (as with any gauge theory), (2.14) implies that $m_{\mu\lambda}$ has an expansion

$$m_{\mu\lambda}=D_{\mu\lambda}{}^{\rho\sigma}\gamma_{\rho\sigma} \qquad (4.9)$$

for some $D_{\mu\lambda}{}^{\rho\sigma}$. The definition (2.15) then implies that

$$M_\mu=-\tfrac{1}{4}K_\mu{}^{\kappa\lambda}\gamma_{\kappa\lambda} \qquad (4.10)$$

for some $K_\mu{}^{\kappa\lambda}$. If we define a new nonsymmetric connection in the tangent space

$$\overline{\Gamma}_{\mu\kappa}{}^{\lambda}=\Gamma_{\mu\kappa}{}^{\lambda}-K_{\mu\kappa}{}^{\lambda}\ , \qquad (4.11)$$

one may easily verify that (2.13) becomes

$$\overline{\nabla}_\mu\gamma_\kappa=0\ , \qquad (4.12)$$

where $\overline{\nabla}_\mu$ is the covariant derivative formed from $\Phi_{\mu a}{}^{b}$ and $\overline{\Gamma}_{\mu\kappa}{}^{\lambda}$. This implies

$$\overline{\nabla}_\mu g_{\kappa\lambda}=0\ . \qquad (4.13)$$

The quantity M_μ in the spin connection can now be given in terms of torsion in the event connection. $K_{\mu\kappa}{}^{\lambda}$ is called the contorsion tensor. $S_{\mu\kappa}{}^{\lambda}\equiv K_{[\mu\kappa]}{}^{\lambda}$ is called the torsion. The contorsion can be expressed in terms of the torsion by[9]

$$K_{\mu\lambda}{}^{\kappa}=S_{\mu\lambda}{}^{\kappa}-S_{\lambda}{}^{\kappa}{}_{\mu}+S^{\kappa}{}_{\mu\lambda}\ . \qquad (4.14)$$

This nonsymmetric metric connection $\overline{\Gamma}_{\mu\kappa}{}^{\lambda}$ and its relation to Dirac spinors has been examined by Hehl and Datta.[23] They apply the formalism to the Einstein action, and find that the torsion gives rise to contact spin-spin interactions and an effective nonlinear spinor equation of the Heisenberg type.[24] Leutwyler did a similar analysis totally in the context of the spin connection.[14] The results of Leutwyler[14] and Hehl and Datta[23] can be applied here. In the low-momentum limit of the M_μ field, the kinetic term of the field [the first of (4.1)] is ignored. Taking also the flat space-time limit, a form of the Heisenberg nonlinear spinor equation is obtained by solving (4.1) in this approximation for M_μ and substituting into (3.5). For higher-momentum transfer, the interaction propagates through the M_μ field removing the contact nature of the spin-spin interaction.

This gives interpretation to all parts of the theory. Half of the spin connection is gravity; the other half is torsion which is associated with spin interactions. The presence of the torsion is a bit unexpected. I began by putting a connection on the tangent space which was explicitly chosen to be symmetric, the torsion chosen to be zero. Even so, the torsion reappeared in the end.

V. CONCLUSION

In this paper, I have attempted to merge two traditional points of view on the spin connection into a Yang-Mills formulation of gravitation. The first point of view is that of Schrödinger, Laurent, and Loos, all of whom studied the gauged spin geometry as the carrier of gravitation.[9-11] The second point of view is that of Leutwyler, who used the spin connection in a true Yang-Mills sense but gave it an interpretation in terms of meson fields. I have amalgamated these two approaches. The resulting theory, as shown in the previous sections, varies slightly from Einstein's in its field equations. A brief analysis of these equations showed them to be reasonable when applied to the simplest observational tests. A complete analysis in the parametrized post-Newtonian formalism, though, is needed.[25]

There are other theories in the literature going under the title of Yang-Mills theories of gravitation. The first work in this direction was done by Loos and Treat.[5] They considered the free Yang-Mills equation for the spin connection. In my notation, they chose (2.13) and (3.4) with μ, M_μ, and j_μ zero. They showed that almost all (in the measure theory sense) of the solutions of these equations led to space-time geometries which were Einstein spaces satisfying

$$R_{\mu\lambda}=\Lambda g_{\mu\lambda} \text{ and } \Lambda \text{ constant.}$$

These are vacuum solutions for Einstein's theory with cosmological constant.

Carmeli's theory is also based on spinors, though he uses two spinors, not Dirac spinors; the gauge group is still SL(2, C) of course. He makes contact with the Newman-Penrose formalism.[26] His kinematics is essentially the same as mine. The difference is in the dynamical statement. His construction is as follows: Form the three mixed quantities

$$F^{W}{}_{\nu\mu} \equiv \tfrac{1}{4} C_{\nu\mu}{}^{\rho\sigma} \gamma_{\rho\sigma} \ , \qquad (5.1)$$

$$J_{\nu\mu} \equiv \tfrac{1}{2} g_{[\nu[\lambda}(T_{\kappa]\mu]} - \tfrac{1}{3} g_{\kappa]\mu]}\, T^{\rho}{}_{\rho}) \ , \qquad (5.2)$$

and

$$F^{J}{}_{\nu\mu} \equiv F^{W}{}_{\nu\mu} + 8\pi G J_{\nu\mu} \qquad (5.3)$$

from the stress-energy tensor $T_{\mu\lambda}$ and the Weyl curvature $C_{\mu\kappa\lambda}{}^{\rho}$. Carmeli's Lagrangian is

$$\mathfrak{L} = \tfrac{1}{2} \epsilon^{\kappa\lambda\mu\rho} \operatorname{tr} F^{J}{}_{\kappa\lambda} (F^{J}{}_{\mu\rho} - 2\partial_{[\mu}\Phi_{\rho]} + [\Phi_{\mu}, \Phi_{\rho}]) \ . \qquad (5.4)$$

He varies $F^{J}{}_{\kappa\lambda}$ and Φ_{μ} independently to give field equations

$$\delta F^{J}{}_{\kappa\lambda}: \ F^{J}{}_{\kappa\lambda} = 2\partial_{[\kappa}\Phi_{\lambda]} - [\Phi_{\kappa}, \Phi_{\lambda}] \qquad (5.5)$$

and

$$\delta\Phi_{\mu}: \ \nabla_{[\mu} F^{J}{}_{\kappa\lambda]} = 0 \ . \qquad (5.6)$$

The first tells us that $F^{J}{}_{\kappa\lambda}$ equals the spin curvature $\phi_{\mu\kappa}$. The second is just the Bianchi identity (2.10a). The conjecture is that Einstein's equations

$$G_{\mu\lambda} = -8\pi G T_{\mu\lambda} \qquad (5.7)$$

satisfy (5.5). Rewrite (5.7) as

$$R_{\mu\lambda} = -8\pi G (T_{\mu\lambda} - \tfrac{1}{2} G_{\mu\lambda} T^{\rho}{}_{\rho}) \ . \qquad (5.8)$$

Use $M_{\mu} = 0$, $\phi_{\mu\kappa} = F^{J}{}_{\mu\kappa}$, and (5.3) in (2.14) along with the expansion of $R_{\mu\kappa\lambda}{}^{\rho}$ by (4.6). Substituting the definitions of $F^{W}{}_{\mu\kappa}$ (5.1) and $J_{\mu\kappa}$ (5.2) gives

$$(\tfrac{1}{4} C_{\nu\mu\lambda\kappa} - \tfrac{1}{2} g_{[\nu[\lambda} R_{\mu]\kappa]} + \tfrac{1}{12} g_{[\nu[\lambda} g_{\mu]\kappa]} R) \gamma^{\lambda\kappa} = \tfrac{1}{4} C_{\nu\mu}{}^{\rho\sigma} \gamma_{\rho\sigma} + 4\pi G [g_{[\nu[\lambda}(T_{\mu]\kappa]} - \tfrac{1}{2} g_{\mu]\kappa]}\, T^{\rho}{}_{\rho}) \gamma^{\lambda\kappa} \ . \qquad (5.9)$$

Showing that (5.8) satisfies (5.9) is straightforward. Contracting (5.9), one can show it to be the only solution. Carmeli's theory reduces to Einstein's theory exactly.

It is stretching the term quite a bit to call Carmeli's theory a Yang-Mills theory. The Lagrangian (5.4) is not the Yang-Mills Lagrangian and the field equations are not in Yang-Mills form. Because of this, one would have difficulty meshing this theory with Yang-Mills–type theories (e.g., Weinberg-Salam style theories of weak and electromagnetic interactions[4]). The source fields are included in this theory in an entirely passive way through $T_{\mu\lambda}$. When we want to write down a complete set of field equations for the other fields in the theory there are difficulties. Add the Dirac Lagrangian (3.2) to (5.4). The field equation (5.6) becomes

$$\delta\Phi_{\mu}: \ \epsilon^{\rho\mu\kappa\lambda} \nabla_{\mu} F^{J}{}_{\kappa\lambda} = j^{\rho} \ , \qquad (5.10)$$

where j^{ρ} is the spin current from (4.1). Eq. (5.5) is unchanged. The Dirac equation (3.5) picks up a complicated source from the $\overline{\psi}$ dependence of $F^{J}{}_{\kappa\lambda}$ through $T_{\mu\lambda}$ in (5.2) and (5.3). Equations (5.10) and (5.5) imply that $j^{\rho} = 0$. This implies that $\psi = 0$, which is physically unacceptable. This indicates serious problems in using Carmeli's theory in a full scale unification with other gauge theories.

Yang's theory[3] when completed with matter sources leads to serious difficulties if done in a straightforward manner. A discussion of these difficulties can be found in a previous article.[20,27] Yang's theory, though fraught with these difficulties, was very much along the line of the philosophy of this paper, to apply strictly the usual Yang-Mills dynamics to the gravitational field.

Camenzind has taken a different point of view from mine in completing Yang's free gravitational equations.[6] My idea was to complete the Yang-Mills variational principle by adding matter sources.[20] Camenzind has made a completion of the field equations themselves directly without a variational principle. Unlike Carmeli and I, who work in the spin spaces of space-time, Camenzind works back in the tangent space with the $SO_0(3,1)$ group.[28] He writes down a Yang-Mills field equation for $SO_0(3,1)$ in the tranditional form using the event connection and curvature

$$\nabla_{\mu} R^{\mu\kappa\rho\sigma} = 8\pi G J^{\kappa\rho\sigma} \ . \qquad (5.11)$$

$J^{\kappa\rho\sigma}$ is some matter current source to gravity, not to be confused with the $J^{\mu\kappa\lambda}$ of (3.7). Camenzind chooses

$$J^{\kappa\rho\sigma} \equiv 2\nabla^{[\rho} T^{\sigma]\kappa} - g^{\kappa[\sigma} \nabla^{\rho]} T^{\lambda}{}_{\lambda} \ , \qquad (5.12)$$

where $T_{\mu\kappa}$ is the matter stress-energy tensor. As with Carmeli's theory (5.11) has Einstein's equations with sources as a solution. To see this, use the once-contracted Bianchi identity [from Eq. (2.11b)] to rewrite (5.11) as we did for (4.5). The equation becomes

$$2\nabla^{[\rho} R^{\sigma]\kappa} = 16\pi G \nabla^{[\rho} T^{\sigma]\kappa} - 8\pi G g^{\kappa[\sigma} \nabla^{\rho]} T^{\lambda}{}_{\lambda} \ . \qquad (5.13)$$

Equation (5.8) satisfies (5.13) by straightforward calculation. Unlike Carmeli's theory, there are other solutions to (5.11) besides (5.8). Camenzind places great importance in these additional solutions to avoid physical singularities which are a hallmark of Einstein's theory.

The main objection to Camenzind's theory is one's inability to put it into a variational form. The usual Yang-Mills current comes from the functional dependence of the interaction Lagrangian on the gauge potential. Currents found by variation will not have the form of (5.12). Sources with nonderivative metric coupling[29] are examples: the variational current is zero even when the stress-energy $T_{\mu\lambda}$ (and thereby $J^{\rho\mu\kappa}$) is nonzero.

Can one unify this approach with other gauge theories and still preserve quantum properties? With no variational principle available, this is highly doubtful. This completes the review of previous Yang-Mills approaches.

The theory given in this paper does not suffer from any of the above difficulties. Also, it is the theory most like that of Yang and Mills which is at the same time observationally viable. It has the following properties:

(a) It is physically acceptable.

(b) It has Yang-Mills-type field equations.

(c) It follows from a variational principle.

(d) The variational principle is of a Yang-Mills type.

(e) It unifies in a straightforward way with other gauge theories.

Yang's theory with variational completion does not have property (a). Carmeli's does not have properties (b) and (d). Camenzind's does not have properties (c) and (d). It is also questionable that either will have property (e). Having property (c) and/or (d) should not be underestimated when one plans to quantize. Even a quantum field approach not directly using a Lagrangian will often rely upon it for the definition of such quantities as conjugate momentum and for structural information. Also, property (d) gives us the hope that some of the quantum properties of the usual Yang-Mills theories might influence this theory of gravity. Because the gauged spin theory of gravity has all of properties (a) through (d) do we get property (e). I shall demonstrate.

Consider a trivial unification. Take any of the unified theories of weak and electromagnetic interactions with gauge group $\mathcal{G}$. The trivial unification, in both the group sense and in difficulty, is to a theory with gauge group $\mathcal{G}\otimes \mathrm{SL}(2,C)$. Add $\mathcal{L}_\Phi(3.1)$ to the Lagrangian of this chosen theory and make all gauge-covariant derivatives gauged spin covariant also. Vary Φ_μ and $g_{\mu\lambda}$ along with the other variables of the theory to get a complete set of equations. This is the trivial unification. A nontrivial unification of $\mathcal{G}$ and $\mathrm{SL}(2,C)$, though, would be much more tantalizing.

APPENDIX

Here I will prove the kinematical relations (2.13), (2.14), and (2.16) of Sec. II. To prove them, I need to develop a bit more on the structure of the Dirac spin space. As usual, one generates the span of the valence $[\begin{smallmatrix}1\\1\end{smallmatrix}]$ spinors using a complete set of Dirac matrices. First one introduces the 16 matrices $\{I,\gamma_\mu,\gamma_{\mu\kappa},\gamma_\mu\gamma,\gamma\}$, where we define $\gamma_{\mu\kappa}$ and γ by

$$\gamma_{\mu\kappa}\equiv\gamma_{[\mu}\gamma_{\kappa]} \tag{A1}$$

and

$$\gamma\equiv\frac{g^{1/2}}{4!}\epsilon^{\mu\kappa\lambda\rho}\gamma_\mu\gamma_\kappa\gamma_\lambda\gamma_\rho\ . \tag{A2}$$

The $\epsilon^{\mu\kappa\lambda\rho}$ is the alternating Levi-Civita symbol.[8] Note that γ is often called γ_5 in the literature. The set $\{I,\gamma_\mu,\gamma\}$ is self-adjoint with respect to the spin metric $g_{a\dot b}$ while $\{\gamma_{\mu\kappa},\gamma_\mu\gamma\}$ is anti-self-adjoint.

The Dirac matrices satisfy the following commutation relations:

$$\begin{aligned}
&[\gamma_\kappa,\gamma_\lambda]=2\gamma_{\kappa\lambda}\ ,\\
&[\gamma_\kappa,\gamma_{\lambda\mu}]=2g_{\kappa\lambda}\gamma_\mu-2g_{\kappa\mu}\gamma_\lambda\ ,\\
&[\gamma_\kappa,\gamma_\lambda\gamma]=2g_{\kappa\lambda}\gamma\ ,\\
&[\gamma_\kappa,\gamma]=2\gamma_\kappa\gamma\ ,\\
&[\gamma_{\kappa\lambda},\gamma_{\mu\rho}]=2g_{\mu\lambda}\gamma_{\kappa\rho}-2g_{\mu\kappa}\gamma_{\lambda\rho}\\
&\qquad\qquad-2g_{\rho\lambda}\gamma_{\kappa\mu}+2g_{\rho\kappa}\gamma_{\lambda\mu}\ ,\\
&[\gamma_{\kappa\lambda},\gamma_\mu\gamma]=2g_{\mu\lambda}\gamma_\kappa\gamma-2g_{\mu\kappa}\gamma_\lambda\gamma\ ,\\
&[\gamma_{\kappa\lambda},\gamma]=0\ ,\\
&[\gamma_\kappa\gamma,\gamma_\lambda\gamma]=2\gamma_{\kappa\lambda}\ ,\\
&[\gamma_\kappa\gamma,\gamma]=-2\gamma_\kappa\ ,\\
&[\gamma,\gamma]=0\ .
\end{aligned} \tag{A3}$$

They also satisfy the following anticommutation relations:

$$\begin{aligned}
&\{\gamma_\mu,\gamma_\kappa\}=2g_{\mu\kappa}I\ ,\\
&\{\gamma_\mu,\gamma_{\kappa\lambda}\}=2g^{1/2}\epsilon_{\mu\kappa\lambda\rho}\gamma^\rho\gamma\ ,\\
&\{\gamma_\mu,\gamma_\kappa\gamma\}=2g^{1/2}\epsilon_{\mu\kappa\lambda\rho}\gamma^{\lambda\rho}\ ,\\
&\{\gamma_\mu,\gamma\}=0\ ,\\
&\{\gamma_{\mu\kappa},\gamma_{\lambda\rho}\}=2g^{1/2}\epsilon_{\mu\kappa\lambda\rho}\gamma-2g_{\mu\lambda}g_{\kappa\rho}I+2g_{\kappa\lambda}g_{\mu\rho}I\ ,\\
&\{\gamma_{\mu\kappa},\gamma_\rho\gamma\}=2g^{1/2}\epsilon_{\mu\kappa\lambda\rho}\gamma^\rho\ ,\\
&\{\gamma_{\mu\kappa},\gamma\}=2g^{1/2}\epsilon_{\mu\kappa\lambda\rho}\gamma^{\lambda\rho}\ ,\\
&\{\gamma_\mu\gamma,\gamma_\kappa\gamma\}=2g_{\mu\kappa}I\ ,\\
&\{\gamma_\mu\gamma,\gamma\}=0\ ,\\
&\{\gamma,\gamma\}=-2\ .
\end{aligned} \tag{A4}$$

These are, of course, just the usual flat Minkowski space-time results taken over into curved space-time. All the usual algebraic properties of the Dirac matrices carry over. For example, all of the 16 matrices are trace-free except for I.

Since these 16 matrices span the space of valence $[\begin{smallmatrix}1\\1\end{smallmatrix}]$ spinors, we can expand a spinor such as $A_a{}^b$ in terms of them:

$$A=aI+b^\mu\gamma_\mu+c^{\mu\lambda}\gamma_{\mu\lambda}+d^\mu\gamma_\mu\gamma+h\gamma\ . \tag{A5}$$

This is called the Clifford expansion. Using (A4), one can invert (A5) to give

$$
\begin{aligned}
a &= \tfrac{1}{4}\mathrm{tr}A \ , \\
b^{\mu} &= \tfrac{1}{4}\mathrm{tr}(A\gamma^{\mu}) \ , \\
c^{\mu\lambda} &= -\tfrac{1}{16}\mathrm{tr}(A\gamma^{\mu\lambda}) \ , \\
d^{\mu} &= \tfrac{1}{4}\mathrm{tr}(A\gamma^{\mu}\gamma) \ , \\
f &= -\tfrac{1}{4}\mathrm{tr}(A\gamma) \ .
\end{aligned}
\tag{A6}
$$

With this bit of algebra behind us, we are in a position to prove the needed results.

Take the covariant derivative ∇_{λ} of (2.13) and antisymmetrize in λ and μ. Using the Liebnitz rule of derivatives and the Ricci identities (2.9) gives

$$
-\tfrac{1}{2}R_{\lambda\mu\kappa}{}^{\rho}\gamma_{\rho} - \tfrac{1}{2}[\phi_{\lambda\mu}, \lambda_{\kappa}] = [\nabla_{[\lambda}M_{\mu]}, \gamma_{\kappa}] + [M_{[\mu}, \nabla_{\lambda]}\gamma_{\kappa}] \ . \tag{A7}
$$

Using (2.13) in the last term, the Jacobi identity of the commutator, and the definition (2.15) gives

$$
-\tfrac{1}{2}R_{\lambda\mu\kappa}{}^{\rho}\gamma_{\rho} - \tfrac{1}{2}[\phi_{\lambda\mu} + m_{\lambda\mu}, \gamma_{\kappa}] = 0 \ . \tag{A8}
$$

By examining the commutation relations (A3) one can see that the Clifford expansion of $\phi_{\lambda\mu} + m_{\lambda\mu}$ can only have $\gamma_{\mu\kappa}$ and I terms:

$$
\phi_{\lambda\mu} + m_{\lambda\mu} = \theta_{\lambda\mu}{}^{\rho\sigma}\gamma_{\rho\sigma} + f_{\lambda\mu}I \ . \tag{A9}
$$

Substituting (A9) into (A8) and using the commutators gives

$$
\theta_{\lambda\mu}{}^{\rho\sigma} = \tfrac{1}{4}R_{\lambda\mu}{}^{\rho\sigma} \ . \tag{A10}
$$

Choice of the holonomy group to be SL(2, C) restricts $f_{\lambda\mu}$ to be zero. This gives us (2.14).

The proof of (2.16) also uses (2.13) and the Clifford expansion. Writing out the covariant derivative in (2.13) gives

$$
\partial_{\mu}\gamma_{\kappa} - \Gamma_{\mu\kappa}{}^{\lambda}\gamma_{\lambda} - [\Phi_{\mu} + M_{\mu}, \gamma_{\kappa}] = 0 \ . \tag{A11}
$$

Expand $\Phi_{\mu} + M_{\mu}$ in a Clifford expansion:

$$
\Phi_{\mu} + M_{\mu} = a_{\mu}I + b_{\mu}{}^{\kappa}\gamma_{\kappa} + c_{\mu}{}^{\kappa\lambda}\gamma_{\kappa\lambda} + d_{\mu}{}^{\kappa}\gamma_{\kappa}\gamma + h_{\mu}\gamma \ . \tag{A12}
$$

Use (A3) to evaluate the commutators term by term. Expand $\partial_{\mu}\gamma_{\kappa}$ also in a Clifford expansion using (A6) and (A7):

$$
\partial_{\mu}\gamma_{\kappa} = \tfrac{1}{4}\mathrm{tr}(\gamma^{\rho}\partial_{\mu}\gamma_{\kappa})\gamma_{\rho} - \tfrac{1}{16}\mathrm{tr}(\gamma^{\rho\sigma}\partial_{\mu}\gamma_{\kappa})\gamma_{\rho\sigma} + \tfrac{1}{4}\mathrm{tr}(\gamma^{\rho}\gamma\partial_{\mu}\gamma_{\kappa})\gamma_{\rho}\gamma - \tfrac{1}{4}\mathrm{tr}(\gamma\partial_{\mu}\gamma_{\kappa})\gamma \ . \tag{A13}
$$

Equating the coefficients of the 16 matrices gives a set of tensor equations which can be solved for $\{b_{\mu}{}^{\kappa}, c_{\mu}{}^{\kappa\lambda}, d_{\mu}{}^{\kappa}, h_{\mu}\}$. This gives us the result (2.16) except for the extra $a_{\mu}I$ term. Using this in (2.10) and comparing to (A9) gives

$$
f_{\lambda\mu} = \partial_{[\lambda}a_{\mu]} \ . \tag{A14}
$$

Our choosing $f_{\lambda\mu} = 0$ implies that a_{μ} can be easily transformed away by a change of frame in the spin space.

We could have kept this electromagnetic part $\{f_{\mu\lambda}, a_{\mu}\}$ of the gauged spin connection as did Schrödinger.[10] The spin gauge holonomy group is then SL(2, C) $\otimes$ U(1). Since I have had no need of it, I have suppressed it.

Finally we are in a position to prove (2.12) if and only if we have (2.13). Differentiating the Dirac relation (2.5) gives

$$
(\nabla_{\mu}\gamma_{(\kappa})\gamma_{\lambda)} + \gamma_{(\kappa|}\nabla_{\mu}|\gamma_{\lambda)} = \nabla_{\mu}g_{\kappa\lambda}I \ . \tag{A15}
$$

It is a simple matter to substitute (2.13) into (A15) to show the left-hand side to be zero. The "only if" part is a bit more tedious.

It is a well-known result that in Minkowski space-time, the metric $\eta_{\mu\kappa}$ determines the Dirac matrices up to a similarity transformation. This is also true in curved space-time as can be seen from (2.5). Choose normal coordinates at some point x so that $g_{\mu\lambda}$ is in Minkowski form at x. Choose γ_{κ} in the standard flat space-time form at x:

$$
\gamma_0 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \ \gamma_1 = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & -1 & & \\ -1 & & & \end{pmatrix},
$$

$$
\gamma_2 = \begin{pmatrix} & & & i \\ & & i & \\ & i & & \\ -i & & & \end{pmatrix}, \ \gamma_3 = \begin{pmatrix} & & 1 & \\ & & & -1 \\ -1 & & & \\ & 1 & & \end{pmatrix}. \tag{A16}
$$

Do this at each point x of space-time. The continuity of $g_{\mu\lambda}(x)$ ensures that $\gamma_{\mu}(x)$ can be chosen this way and be continuous also.

We are given some metric in the tangent space which satisfies the metric constraint (2.12). Construct a set of γ_{μ} in the way indicated above. Construct the Cristoffel connection. Choose any $M_{\mu a}{}^{b}$. From these define a $\Phi_{\mu a}{}^{b}$ using Eq. (2.16). Substitute this form into (2.13) [or (A11)]. One finds that after a bit of tedious algebra using the commutation relations (A3) that the equation reduces to

$$
\Gamma_{\mu(\kappa}{}^{\rho}g_{\lambda)\rho} = -\tfrac{1}{4}\mathrm{tr}(\gamma_{(\lambda|}\partial_{\mu}|\gamma_{\kappa)}) \ . \tag{A17}
$$

Permuting under the trace, the equation can be rewritten as

$$
\Gamma_{\mu(\kappa}{}^{\rho}g_{\lambda)\rho} = -\tfrac{1}{8}\mathrm{tr}(\partial_{\mu}g_{\kappa\lambda}I) = -\tfrac{1}{2}\partial_{\mu}g_{\kappa\lambda} \ . \tag{A18}
$$

But this follows from the Christoffel form of the connection,[18] thus completing the proof.

The last technical detail that must be considered is Eq. (3.14). I have taken the Dirac γ matrices as dynamical functionals of the metric. The Dirac relation (2.5) tells us that we could have done the reverse easily enough, but is (3.14) feasible? Let us consider arbitrary variations in γ_μ. Take the Clifford expansion of $\delta\gamma_\mu$,

$$\delta\gamma_\mu \equiv a_\mu I + b_\mu{}^\kappa\gamma_\kappa + c_\mu{}^{\kappa\lambda}\gamma_{\kappa\lambda} + d_\mu{}^\kappa\gamma_\kappa\gamma + h_\mu\gamma \ , \tag{A19}$$

for infinitesimal coefficients. Substitute this into the variation of the Dirac relation (2.5) which is

$$\delta\gamma_{(\kappa}\gamma_{\lambda)} + \gamma_{(\kappa}\delta\gamma_{\lambda)} = \delta g_{\kappa\lambda} I \ . \tag{A20}$$

One finds that

$$\delta\gamma_\mu = \tfrac{1}{2}\delta g_{\mu\kappa}\gamma^\kappa + \epsilon_{\mu\kappa}\gamma^\kappa + h_\mu\gamma \ ,$$
$$\epsilon_{\mu\kappa} \equiv b_{[\mu\kappa]} \ . \tag{A21}$$

The first term is just (3.24). Thus the variations of γ_μ can be classified into independent categories: (1) those of type (3.14) and (2) those of the type given by the last two terms of (A21).

When we let the spin connection be zero and the metric be flat and of Minkowski form, the theory must physically reduce to the usual Dirac theory of a single electron. This requirement prohibits us from making variations of the second type. To see this, do these variations in the action (3.3). A complete set of variations would be

$$\{\delta\Phi_\mu, \delta\psi, \delta\bar\psi, \delta g_{\mu\kappa}, \delta\gamma_\mu = \epsilon_{\mu\kappa}\gamma^\kappa + h_\mu\gamma\} \ .$$

The first four give Eqs. (3.4), (3.5) and its complex conjugate, and (3.9). The last in the flat space-time limit gives

$$\bar\psi\gamma_{[\mu}\partial_{\kappa]}\psi - (\partial_{[\kappa}\bar\psi)\gamma_{\mu]}\psi = 0 \tag{A22}$$

and

$$\bar\psi\gamma\partial_\kappa\psi - (\partial_\kappa\bar\psi)\gamma\psi = 0 \ . \tag{A23}$$

These equations severely limit the solution set of the Dirac equation and do so in a way which strongly depends on the representation of the Dirac matrices. This is totally unacceptable. Thus to preserve the proper flat space-time limit, we must restrict the variations of γ_μ to be of the form of (3.14).

Another way to see this is to note that the freedom

$$\delta\gamma_\kappa = \epsilon_{\kappa\lambda}\gamma^\lambda + h_\kappa\gamma \tag{A24}$$

for fixed metric $g_{\mu\lambda}$ corresponds to a freedom of choosing the representation of γ_μ. I chose (A16), but any other *fixed* choice is acceptable. This freedom arises from the lack of a natural map from the spin space to the tangent space. This is a kinematical freedom, not a dynamical one; it should not be varied in the action. Rather, this aspect of the γ_μ should be fixed in the beginning and carried through the problem unchanged. Equation (3.14) identifies the only dynamic freedom in γ_μ.

Note added in proof. By using the usual covariant quantization and background field methods, I have been able to show that this theory is n-loop renormalizable. This result has been checked to one loop using the 't Hooft and Veltman algorithm in curved space-time. It does this by avoiding the dipole ghost and nonunitarity problems of the renormalizable (R^2+R)-type quantum gravity theories to which this one is closely related.

ACKNOWLEDGMENTS

I would like to thank R. P. Treat for numerous discussions on the spin connection and to C. DeWitt-Morette for constant encouragement during the early stages of this research. I very much appreciate the hospitality shown me by D. W. Sciama and the theory group of the Department of Astrophysics, Oxford University and the many helpful discussions.

---

*Work supported in part by a NATO postdoctoral fellowship at the Department of Astrophysics, Oxford University.

†Permanent address.

[1]That Einstein's theory is a local gauge theory was first noted by R. Utiyama, Phys. Rev. 101, 1597 (1956).

[2]C. N. Yang and R. L. Mills, Phys. Rev. 96, 191 (1954).

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

[4]S. Weinberg, Rev. Mod. Phys. 46, 255 (1974); E. S. Abers and B. W. Lee, Phys. Rep. 9C, 1 (1973); J. Bernstein, Rev. Mod. Phys. 46, 7 (1974), and references therein.

[5]H. G. Loos and R. P. Treat, Phys. Lett. 26A, 91 (1967).

[6]M. Camenzind, J. Math. Phys. 16, 1023 (1975); M. Camenzind, Gen. Rel. Gravit. 8, 103 (1977); M. Camenzind and M. A. Camenzind, Instit. Theor. Phys., Univ. of Hamburg report (unpublished).

[7]M. Carmeli, Phys. Rev. D 14, 1727 (1976); Nucl. Phys. B38, 621 (1972); J. Math. Phys. 11, 2728 (1970); M. Carmeli and S. I. Fickler, Phys. Rev. D 5, 290 (1972).

[8]H. G. Loos, J. Math. Phys. 8, 2114 (1967). I will speak of the gauge holonomy group of a theory as the maximal holonomy group of any solution of the field equations of the theory.

[9]J. A. Schouten, *Ricci-Calculus* (Springer, Berlin, 1954). I follow this reference most closely for conventions. My Ricci tensor is given by

$R_{\mu\kappa} = R_{\gamma\mu\kappa}{}^{\lambda}$.

[10]E. Schrödinger, Sitzungsber. Preuss. Akad. Wiss. Phys.-Math. KL XI, 105 (1932).

[11]B. E. Laurent, Ark. Fys. 16, 263 (1959).

[12]H. G. Loos, Ann. Phys. (N.Y.) 25, 91 (1963).

[13]R. P. Treat, J. Math. Phys. 11, 2187 (1970).

[14]H. Leutwyler, Nuovo Cimento 26, 1066 (1962).

[15]These are some of the more important references to the spin connection besides Refs. 5 and 10–14: H. Weyl, Z. Phys. 56, 330 (1929); V. Fock, *ibid.* 57, 261 (1929); J. A. Schouten, Stud. Appl. Math. 10, 239 (1931); V. Bargmann, Sitzungsber. Preuss. Akad. Wiss. Phys.-Math. KL XI, 346 (1932); L. Infeld and B. L. van der Waerden, *ibid.* XII, 380 (1933); F. J. Belinfante, Physica (Utrecht) 7, 305 (1940); D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. 29, 465 (1952); J. G. Fletcher, Nuovo Cimento 8, 451 (1958); O. Klein, Ark. Fys. 17, 517 (1960); T. Kimura, Prog. Theor. Phys. 24, 386 (1960); J. R. Klauder, H. Leutwyler, and J. Schaer, Nuovo Cimento 24, 389 (1962); A. Peres. *ibid.* 24, 389 (1962); H. G. Loos, *ibid.* 30, 901 (1963); A. Brown, J. Math. Ann. Appl. 25, 537 (1969); M. Novello, J. Math. Phys. 12, 1039 (1971); F. W. Hehl and B. K. Datta, *ibid.* 12, 1334 (1971); S. K. Wong, Int. J. Theor. Phys. 5, 221 (1972); C. P. Luehr and M. Rosenbaum, J. Math. Phys. 15, 1120 (1974); other references can be found in these.

[16]The Hermitian condition of $g_{a\dot{b}}$ is

$$(g_{a\dot{b}})^* = g_{\dot{a}b} = g_{\dot{b}a},$$

where the asterisk denotes complex conjugation.

[17]A coordinate frame is also known as a holonomic frame (or holonomic coordinates in Ref. 9).

[18]

$$\Gamma_{\mu\kappa}{}^{\lambda} = \tfrac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\kappa\rho} + \partial_\kappa g_{\mu\rho} - \partial_\rho g_{\mu\kappa}).$$

[19]I am using units with $\hbar = c = 1$.

[20]E. E. Fairchild, Jr., Phys. Rev. D 14, 384 (1976).

[21]G. B. Debney, E. E. Fairchild, Jr. and S. T. C. Siklos, unpublished result. This result was proved by putting (4.7) and (4.8) in a null tetrad and using the freedom of the tetrad to put R_μ in canonical form.

[22]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972). My convention for the Ricci tensor as well as my choice of metric signature is the negative of Weinberg's.

[23]F. W. Hehl and B. K. Datta, J. Math. Phys. 12, 1334 (1971). A review of the usual torsion version of Einstein's theory (the Sciama-Kibble type modeled on a metric nonsymmetric Riemann-Cartan geometry) is given by F. W. Hehl *et al.*, Rev. Mod. Phys. 48, 393 (1976).

[24]W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Wiley, London, 1966).

[25]C. M. Will, Phys. Today, 25 (No. 10), 23 (1972); Astrophys. J. 163, 611 (1971); 185, 31 (1973); C. M. Will and K. S. Thorne, *ibid.* 163, 595 (1971); C. M. Will and K. Nordtvedt, Jr., *ibid.* 177, 757 (1972); 177, 775 (1972); W.-T. Ni, *ibid.* 176, 769 (1972).

[26]E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962). See also F. A. E. Pirani, in *Lectures on General Relativity,* edited by Deser and Ford (Prentice-Hall, Englewood Cliffs, New Jersey, 1965).

[27]An error exists in Ref. 20 in Eq. (29) invalidating the lemma of the second section which proves that the only vacuum solutions to the free gravitational equations in Yang's theory are Einstein spaces. Specifically, the nonphysical solution [Eq. (22)] is still a solution. This makes the difficulties of Yang's theory as summarized there even more serious.

[28]He states his theory in terms of $SO_0(3,1)$ principal bundles and uses a connection form which is equivalent to how I have presented it here; only the notation differs.

[29]A source field has nonderivative coupling with the gravitational field if its interaction Lagrangian depends on the metric algebraically, no connection or curvature dependence. The scalar field, the Dirac field as formulated here, any Yang-Mills field, and the electromagnetic field are examples of gravitational nonderivative coupling. The conformal scalar field exhibits derivative coupling.