Gauge theory of gravitation

Edward E. Fairchild, Jr.
McDonnell Center for the Space Sciences and Department of Physics, Washington University, St. Louis, Missouri 63130
(Received 12 February 1976)

Yang has laid the foundations for a "gauge" type theory of gravitation, reintroducing the Lagrangian previously considered by Weyl and Stephenson. In this paper, I develop the full theory using a variational principle on this Lagrangian with sources. Analysis of the full set of Euler-Lagrange equations shows that Einstein spaces satisfying $R_{\mu\kappa} = \omega g_{\mu\kappa}$, with arbitrary cosmological constant ω, are the only Riemannian vacuum solutions. This rules out the nonphysical, static, spherically symmetric solutions of Pavelle and Thompson; they only considered a subset of the Euler-Lagrange system of equations. The additional equations become important when sources to the gravitational field are considered. The full set of equations with sources have the following properties: The stress-energy tensor must be traceless. Other than a small exceptional class, all matter solutions must be non-Riemannian. The stress-energy tensor is not conserved if torsion is kept. The only Robertson-Walker cosmological solution which is Riemannian is static; all other homogeneous, isotropic cosmologies are non-Riemannian. The equations do not reduce to Poisson's equation for weak, static gravitational fields, thus violating the Newtonian limit. I finish by commenting on the "conceptually superior ... integral formalism" proposed by Yang and used as a foundation for his gauge-type gravitational theory.

INTRODUCTION

Yang has proposed an alternate theory of gravitation to replace Einstein's.[1] This theory was previously proposed in greater detail by Stephenson.[2] The most important difference between this one and Einstein's is that they replace the Lagrangian of Einstein's theory which is linear in the curvature by one quadratic. This Lagrangian is one once proposed by Weyl for his unified field theory of gravitation and electromagnetism.[3] Though the Lagrangian is the same, the resulting field equations are different owing to a difference in constraints.

Thompson[4] and Pavelle[5] have shown that there are nonphysical static, spherically symmetric solutions of the resulting field equations. These solutions either violate the Newtonian limit, or give incorrect values for the perihelion precession of Mercury. To get these field equations, one must vary the connection and metric independently. This is what one does when imposing the Palatini variation on the Einstein Lagrangian.[6] For the Stephenson-Yang case, the connection must be nonmetric and may be nonsymmetric.[7,8] Since the connection and metric are varied independently, one gets two sets of tensor-field equations. Only the first set was noted by Yang. Under the Riemannian constraint of the solution set (i.e., metric-symmetric connections), this first set is equivalent to a set studied by Kilmister and Newman,[9] and Loos,[10] and Loos and Treat.[11] The second set comes from variation of the metric and was not explicitly noted by Yang[12] but developed by Stephenson.[2] Imposing this second set eliminates all of the known nonphysical, static, spherically symmetric solutions. Also, through these second equations, rather than the first, the gravitational field couples to matter sources.

In Sec. I, I will derive the field equations, motivating them as a GL$(4,R)$ gauge theory following Yang.[1] In Sec. II, I will derive the second set of field equations and show that in the Riemannian, matter-free case, the only solutions are Einstein spaces. It is this that rules out the nonphysical solutions. Section III demonstrates what occurs when matter sources are introduced. I will show that the matter stress-energy tensor must be traceless, the Riemannian condition on the connection cannot be kept, and the Newtonian limit is only satisfied in part. In the last section, I will discuss the theory in the light of the theoretical difficulties encountered. Also, I will discuss Yang's "integral formulation"[1] of gauge fields and contrast it with the traditional approach.

I. THE LAGRANGIAN AND THE YANG FIELD EQUATIONS

The idea of gauge fields was first introduced by Yang and Mills[13] who considered local isospin transformations. This leads to the introduction of a gauge-covariant derivative[14]:

$$\nabla_\mu \psi_i = \partial_\mu \psi_i - \Phi_{\mu i}{}^j \psi_j . \qquad (1)$$

This is needed because the partial derivative of ψ_i no longer transforms as an isospin tensor (i.e., not homogeneously) under the local SU(2) isospin transformation:

$$\psi_i \to \psi_{i'} = S_{i'}{}^j(x)\psi_j . \qquad (2)$$

To ensure that $\nabla_\mu\psi_i$ is an isotensor, the gauge linear connection $\Phi_{\mu i}{}^j$ (also called the gauge potentials) must transform via

$$\Phi_{\mu i}{}^{j} \to \Phi_{\mu i'}{}^{j'} = S_{i'}{}^{i}\Phi_{\mu i}{}^{j}S^{-1}{}_{j}{}^{j'} - S_{i'}{}^{i}\partial_{\mu}S^{-1}{}_{i}{}^{j'} . \quad (3)$$

Introducing a matrix notation which suppresses the Latin indices,

$$\Phi_{\mu} \to \Phi'_{\mu} = S\Phi_{\mu}S^{-1} - S\partial_{\mu}S^{-1} . \quad (4)$$

The $\Phi_{\mu i}{}^{j}$ is therefore not an isospin tensor.

Introducing the field quantities $\Phi_{\mu i}{}^{j}$ requires us to specify their field equations. One does this via an action principle. For a Lagrangian which is both a scalar and isoscalar, one needs an isospin tensor formed from $\Phi_{\mu i}{}^{j}$ from which to construct it. The isospin curvature (also called the gauge fields) $\Phi_{\mu\kappa i}{}^{j}$ defined by

$$\phi_{\mu\kappa} = 2\partial_{[\mu}\Phi_{\kappa]} - [\Phi_{\mu}, \Phi_{\kappa}] \quad (5)$$

is the simplest.[15]

The field equations are derived from an action using

$$\mathcal{L}_{\Phi} = \frac{1}{4f^2}\,\phi_{\mu\kappa i}{}^{j}\phi_{\lambda\epsilon j}{}^{i}g^{\mu\lambda}g^{\kappa\epsilon} . \quad (6)$$

as the Lagrangian. The $g^{\mu\lambda}$ is the Minkowski space-time metric, my choice of signature being $(+---)$, and f is the coupling constant of the gauge field. Adding the Lagrangian $\mathcal{L}_{\psi}$ for the isospinor field ψ_i gives a total Lagrangian $\mathcal{L} = \mathcal{L}_{\Phi} + \mathcal{L}_{\psi}$. The interaction is through minimal coupling via the covariant derivative [see Eq. (1)]. These considerations need not be restricted to SU(2), of course, but can be applied to any Lie group in some representation. Electromagnetism from U(1) is the most well-known example. The choice of Eq. (6) as the free Lagrangian for Φ_{μ} is by analogy with electromagnetism. It is the simplest Lagrangian that one can choose.[16] The field equations are

$$\delta\Phi_{\mu i}{}^{j}:\ \frac{1}{f^2}\nabla_{\mu}\phi^{\mu\kappa}{}_{i}{}^{k} = -j^{\kappa}{}_{i}{}^{k} \equiv \frac{\delta S_{\psi}}{\delta\Phi_{\kappa k}{}^{i}} . \quad (7)$$

The $j^{\kappa}{}_{i}{}^{k}$ is the gauge current and S_{ψ} is the action constructed from $\mathcal{L}_{\psi}$.

Yang,[1] and a number of other authors previously,[17-19] have noted the following about the tangent bundle of space-time: If $x^{\mu'}(x^{\kappa})$ is a change of coordinates on space-time, vectors such as v^{κ} transform via

$$v^{\kappa'} = A^{\kappa'}{}_{\mu}(x)v^{\mu},$$
$$A^{\kappa'}{}_{\mu}(x) \equiv \frac{\partial x^{\kappa'}}{\partial x^{\mu}}(x) . \quad (8)$$

The quantity $A^{\kappa'}{}_{\mu}(x)$, at each x, is a 4×4 invertible matrix; therefore, it is an element of the group GL(4,R). If one introduces a gauge-covariant derivative for GL(4,R), as previously outlined, one gets a gauge potential, denoted $\Gamma_{\mu\kappa}{}^{\lambda}$, the linear connection of the tangent bundle. The gauge field is the Riemann-Christoffel curvature of the tangent bundle

$$R_{\mu\kappa\lambda}{}^{\epsilon} \equiv 2\partial_{[\mu}\Gamma_{\kappa]\lambda}{}^{\epsilon} - 2\Gamma_{[\mu|\lambda}{}^{\rho}{}_{|}\Gamma_{\kappa]\rho}{}^{\epsilon} . \quad (9)$$

There are certain subtleties in treating the connection of the tangent bundle as a GL(4,R) gauge potential.[20] I mention only one which is needed in doing the action variations. Consider a general space-time manifold (not necessarily Riemannian) with metric $g_{\mu\kappa}$. The action for any physical system is the integral of the Lagrangian with the invariant measure $d\mu = g^{1/2}dv$. The dv is the usual volume element and g is defined from the metric via $g \equiv -\det(g_{\mu\kappa})$. When the metric is varied as a dynamic quantity, one must vary both the Lagrangian and the measure.

Yang's idea for a gravitation theory is to take the gauge prescription, including choice of dynamics, over entirely and to employ the Lagrangian

$$\mathcal{L}_{\Gamma} = \frac{-1}{4\kappa}\,R_{\mu\kappa\rho}{}^{\sigma}R_{\lambda\epsilon\sigma}{}^{\rho}g^{\mu\lambda}g^{\kappa\epsilon} \quad (10)$$

as the free Lagrangian of gravity. This is in contrast with the Einstein Lagrangian

$$\mathcal{L}_{\Gamma} = -\frac{1}{\kappa}R_{\kappa\mu\lambda}{}^{\kappa}g^{\mu\lambda} \equiv -\frac{1}{\kappa}R . \quad (11)$$

At first glance, this seems an attractive idea. One might easily see this as a first step toward a unified theory of weak, electromagnetic, strong, *and* gravitational interactions.

To vary the action defined by (10) to get the equations proposed by Yang, one must act in strict analogy with the general gauge case. Note these two points: First, the connection $\Gamma_{\mu\kappa}{}^{\lambda}$ is different from the general $\Phi_{\mu i}{}^{j}$ in that the two lower indices are of the same type. Since $\Phi_{\mu i}{}^{j}$ can have no symmetry between μ and i, we should not assume any symmetry of $\Gamma_{\mu\kappa}{}^{\lambda}$ in μ and κ. Second, since the $\Phi_{\mu i}{}^{j}$ has no dynamical relation to the metric $g_{\mu\kappa}$, by analogy $\Gamma_{\mu\kappa}{}^{\lambda}$ and $g_{\mu\kappa}$ must be varied independently. In other words, the connection must be nonsymmetric and nonmetric.[8] It will turn out that choosing the connection symmetric would still give the same equations but would break the analogy.

The Euler-Lagrange equations deduced by varying $\Gamma_{\mu\kappa}{}^{\lambda}$ in the action of (10) are

$$\delta\Gamma_{\mu\kappa}{}^{\lambda}:\ \nabla_{\rho}g^{1/2}R^{\rho\mu}_{\cdot\cdot\lambda}{}^{\kappa} = 0 . \quad (12)$$

The dots under the first two free indices indicate that they have been raised from their defining positions using the metric. Because the metric does not commute with covariant differentiation, I have employed this convention to remember the now critical, hidden metrics. The $g^{1/2}$ comes from the invariant measure.

Let us denote the covariant derivative of the metric by

$$Q_{\mu\kappa\lambda} \equiv -\nabla_\mu g_{\kappa\lambda} \tag{13}$$

and the torsion by

$$S_{\mu\kappa}{}^{\lambda} \equiv \Gamma_{[\mu\kappa]}{}^{\lambda} . \tag{14}$$

If one assumes a Riemannian (i.e., symmetric, metric) connection with $Q_{\mu\kappa\lambda}$ and $S_{\mu\kappa}{}^{\lambda}$ both zero, (12) can be rewritten as

$$\nabla_\rho R_{\kappa\lambda\sigma}{}^{\rho} = 0 . \tag{15}$$

The Bianchi identity in a Riemannian space is

$$\nabla_{[\mu} R_{\kappa\lambda]\sigma}{}^{\rho} = 0 . \tag{16}$$

When contracted on μ and σ and used with Eq. (15) it gives the alternative form of (15):

$$\nabla_{[\kappa} R_{\lambda]\sigma} = 0 . \tag{17}$$

The $R_{\mu\kappa} \equiv R_{\nu\mu\kappa}{}^{\nu}$ is the Ricci tensor.

Yang presents (17) as his alternate to the Einstein free-gravitational-field equations:

$$R_{\mu\kappa} = 0 . \tag{18}$$

He notes that all solutions of (18) are solutions of (17) also, so that the Schwarzchild solution used in the standard gravitational tests of Einstein's theory is preserved.

Note that including torsion during the variation is necessary if one is to keep the strict analogy with the general gauge case. Suppressing it would set apart, in some sense, this gauge field from any other. Requiring $S_{\mu\kappa}{}^{\lambda}=0$ in the variation gives, instead of (12), the field equation

$$\frac{1}{\kappa} \nabla_\rho g^{1/2} R^{\rho(\mu}{}_{\cdot\,\cdot\,\lambda}{}^{\kappa)} = 0 . \tag{19}$$

See the effect of the symmetry of $\Gamma_{\mu\kappa}{}^{\lambda}$ in μ and κ. This equation is weaker than (12). Under $Q_{\mu\kappa\lambda}=0$, it is equivalent to

$$\nabla_\rho R_{\kappa(\lambda\sigma)}{}^{\rho} = 0 . \tag{20}$$

This still gives (17), using the contracted form of (16), but now the symmetry of $R_{\mu\kappa}$ plays an important role in the derivation.

If one were to try to reduce (12) to a form similar to (17), the result would be very complicated and not at all illuminating. One must account for the noncommutativity of ∇_μ and $g_{\mu\kappa}$ and the fact that the Bianchi identity in the presence of torsion is[7]

$$\nabla_{[\mu} R_{\kappa\lambda]\sigma}{}^{\rho} = 2 S_{[\mu\kappa}{}^{\epsilon} R_{\lambda]\epsilon\sigma}{}^{\rho} . \tag{21}$$

II. STATIC SPHERICALLY SYMMETRIC EMPTY-SPACE SOLUTIONS AND THE OTHER SET OF FIELD EQUATIONS

Simply because the Schwarzschild solution satisfies the Yang equations (17) does not imply that the light-bending and perihelion-precession tests will be satisfied as Yang supposes.[1] The Schwarzschild solution is the *unique* static spherically symmetric solution to the source-free Einstein equations (18). Yang's (17) is weaker than (18), and one has no guarantee that the Schwarzschild solution is the proper choice.[21]

Thompson[4] and Pavelle[5] have demonstrated a number of other static spherically symmetric solutions of (17) which are not physical. Either they violate the astronomical tests and/or violate the Newtonian limit. They are given by the line elements

$$ds^2 = (1+MG/r)^{-2}dt^2 - (1+MG/r)^{-2}dr^2 - r^2 d\Omega^2 , \tag{22}$$

$$ds^2 = dt^2 - (1-2MG/r)^{-1}dr^2 - r^2 d\Omega . \tag{23}$$

Compare these to the Schwarzschild solution

$$ds^2 = (1-2MG/r)dt^2 - (1-2MG/r)^{-1}dr^2 - r^2 d\Omega^2 . \tag{24}$$

Pavelle[5] indicates that further restrictions must be placed on the solution set of (17).

There are further restrictions on the solution set. They are given by the second set of field equations, coming from varying the metric. The Euler-Lagrange equations obtained from (10) by varying $g^{\mu\kappa}$ and using the identity $\delta g = -g g_{\mu\kappa}\delta g^{\mu\kappa}$ are

$$\delta g^{\mu\kappa} :\; g^{1/2} H_{\mu\kappa} = 0 , \tag{25}$$

where

$$H_{\mu\kappa} \equiv R_{\mu}{}^{\lambda}{}_{\cdot\,\sigma}{}^{\rho} R_{\kappa\lambda\rho}{}^{\sigma} - \tfrac{1}{4} g_{\mu\kappa} R^{\epsilon\,\lambda}_{\cdot\,\cdot\,\sigma}{}^{\rho} R_{\epsilon\lambda\rho}{}^{\sigma} . \tag{26}$$

The field equation $H_{\mu\kappa}=0$ was developed by Stephenson[2] and must be imposed on the solution set of (17). Note that $H_{\mu\kappa}$ is traceless; $H_{\mu}{}^{\mu}=0$.

One gets a useful alternate form for the case $Q_{\mu\kappa\lambda}=0$ by expressing the curvature in terms of the traceless Weyl conformal curvature, the Ricci curvature, and the scalar curvature[7]:

$$R_{\nu\mu\lambda}{}^{\kappa} = C_{\nu\mu\lambda}{}^{\kappa} - 2 g_{[\nu[\lambda} R_{\mu]\rho]} g^{\rho\kappa} + \tfrac{1}{3} g_{[\nu[\lambda} g_{\mu]\rho]} g^{\rho\kappa} R , \tag{27}$$

with $R \equiv g^{\mu\kappa} R_{\mu\kappa}$. Noting the identity[22]

$$C_{\mu}{}^{\lambda}{}_{\cdot\,\sigma}{}^{\rho} C_{\kappa\lambda\rho}{}^{\sigma} - \tfrac{1}{4} g_{\mu\kappa} C^{\epsilon\,\lambda}_{\cdot\,\cdot\,\sigma}{}^{\rho} C_{\epsilon\lambda\rho}{}^{\rho} = 0 , \tag{28}$$

one gets $H_{\mu\kappa}$ in the form

$$H_{\mu\kappa}=\tfrac{1}{2}g_{\mu\kappa}R_{\sigma\rho}R^{\sigma\rho}_{\cdot\cdot}+\tfrac{5}{3}R_{\mu\kappa}R-2R_{\mu}{}^{\epsilon}_{\cdot}R_{\epsilon\nu}-\tfrac{5}{12}g_{\mu\kappa}R^2$$
$$+C_{\mu\sigma\kappa}{}^{\rho}R_{\rho}{}^{\sigma}_{\cdot}\,. \qquad (29)$$

It is obvious that (18) satisfies both sets of equations. Are there other empty-space solutions? Consider the class of solutions where the Ricci tensor has only a trace part: $R_{\mu\kappa}=\frac{1}{4}Rg_{\mu\kappa}$. One may easily verify that this type satisfies $H_{\mu\kappa}=0$, remembering that $C_{\mu\kappa\lambda}{}^{\epsilon}$ is traceless. The Yang equations (17) imply that R is constant; therefore a class of empty-space solutions are the Einstein spaces given by the requirement

$$R_{\mu\kappa}=\omega g_{\mu\kappa} \qquad (30)$$

for all constants ω. I wish to argue that these are the only empty-space solutions.

On the basis of (17), Loos and Treat[10,11] have shown that other than a small exceptional class of solutions of (17) "of measure zero" the Einstein spaces constitute a complete set. Unfortunately, the static spherically symmetric solutions not satisfying (30) fall in this exceptional class; therefore, one cannot rule out these nonphysical cases by their argument. But we have also the second equation (25). It is strong enough to rule out the exceptional cases (22) and (23). In fact the following lemma is true when the geometry is Riemannian: $R_{\mu\kappa}=\frac{1}{4}Rg_{\mu\kappa}$ *if and only if* $H_{\mu\kappa}=0$.[23] I noted the *only-if* implication above and it is straightforward to show. For the *if* implication, decompose $R_{\mu\kappa}$ into its trace and traceless (denoted $P_{\mu\kappa}$) parts:

$$R_{\mu\kappa}=\tfrac{1}{4}Rg_{\mu\kappa}+P_{\mu\kappa}\,. \qquad (31)$$

Substituting in (29) gives $H_{\mu\kappa}$ as

$$H_{\mu\kappa}=\tfrac{1}{2}g_{\mu\kappa}P_{\lambda\sigma}P^{\lambda\sigma}_{\cdot\cdot}+\tfrac{2}{3}P_{\mu\kappa}R-2P_{\mu}{}^{\sigma}_{\cdot}P_{\sigma\nu}+C_{\mu\sigma\kappa}{}^{\rho}P^{\sigma}_{\cdot\rho}\,. \qquad (32)$$

The burden of proof is to show that $H_{\mu\kappa}=0$ implies $P_{\mu\kappa}=0$. Choose a nonholonomic frame in which both $g_{\mu\kappa}$ and $P_{\mu\kappa}$ are diagonal and $g_{\mu\kappa}$ is in Minkowski form.[24] This is the frame used to define the gravitational invariants[25] which can be taken as the independent components of $C_{\mu\kappa\lambda\epsilon}$ and the diagonal components of $R_{\mu\kappa}$. Equation (32) plus $P_{\mu}{}^{\mu}_{\cdot}=0$ are up to 10 independent nonlinear algebraic equations that the four diagonal components of $P_{\mu\kappa}$ must satisfy. Except when most of the components of $C_{\mu\kappa\lambda\epsilon}$ are zero, the solution set is highly overdetermined and it is easy to show that $P_{\mu\kappa}=0$ is the only solution.

I will outline the proof. Divide the equations into two sets: first, the set $H_{\mu\kappa}=0$ where $\mu=\kappa$ is three independent equations which I will call the diagonal set. They depend only on the scalar curvature R and the gravitational invariants C_{0101} and C_{0202}. Second, the off-diagonal set with $\mu\neq\kappa$ is six equations. They depend only on the gravitational invariants C_{0102}, C_{0103}, C_{0112}, C_{0113}, C_{0203}, and C_{0212}, each equation containing only one. If any one of this second set of invariants is nonzero, its equation implies that two of the diagonal components of $P_{\mu}{}^{\kappa}_{\cdot}$ are equal. The diagonal equations plus $P_{\mu}{}^{\mu}_{\cdot}=0$ are the four equations to solve for three unknowns, which one can easily show has only the solution $P_{\mu\kappa}=0$. When all of the second set of invariants are zero, one has four equations for four unknowns which because of the nonlinearities may have other solutions besides $P_{\mu\kappa}=0$. It is tedious but straightforward to show that there are no others.

Applying this lemma, the Stephenson-Yang equations give the Einstein spaces as the only vacuum solutions. One can solve for the static, spherically symmetric solution of (30) easily following the derivation of the Schwarzschild solution.[26] Its line element is

$$ds^2=\left(1-\frac{2mG}{r}+\frac{\omega r^2}{3}\right)dt$$
$$-\left(1-\frac{2mG}{r}+\frac{\omega r^2}{3}\right)^{-1}dr^2-r^2d\Omega^2\,. \qquad (33)$$

As one expects, the "cosmological constant" ω has effect only at large distances r from the source. It will have noticeably little effect on the standard tests, unless ω is large. This eliminates essentially the problem of the nonphysical solutions.

III. INTRODUCING SOURCES TO THE GRAVITATIONAL FIELD: THE COSMOLOGICAL SOLUTIONS

In introducing the equations, both Stephenson and Yang left open the problem of sources to the gravitational field. Consider a Lagrangian $\mathfrak{L}_M$ for matter and radiation. The sources of equations (12) and (25) follow from the dependence of $\mathfrak{L}_M$ on $\Gamma_{\mu\kappa}{}^{\lambda}$ and $g^{\mu\kappa}$, respectively. Constructing the total Lagrangian $\mathfrak{L}=\mathfrak{L}_\Gamma+\mathfrak{L}_M$, and varying $\Gamma_{\mu\kappa}{}^{\lambda}$ and $g^{\mu\kappa}$ gives the field equations

$$\delta\Gamma_{\mu\kappa}{}^{\lambda}:\quad \nabla_\rho g^{1/2}R^{\rho\mu}_{\cdot\cdot}{}_{\lambda}{}^{\kappa}=\kappa g^{1/2}J^{\mu\kappa}{}_{\lambda}\equiv\kappa\frac{\delta S_M}{\delta\Gamma_{\mu\kappa}{}^{\lambda}}\,, \qquad (34)$$

$$\delta g^{\mu\kappa}:\quad g^{1/2}H_{\mu\kappa}=\kappa g^{1/2}T_{\mu\kappa}\equiv 2\kappa\frac{\delta S_M}{\delta g^{\mu\kappa}}\,. \qquad (35)$$

$T_{\mu\kappa}$ is the stress-energy tensor of relativity; this tensor is well understood. Equation (34) is the equation introduced by Loos to describe spin density in Einstein's theory.[10] $J^{\mu\kappa}{}_{\lambda}$ is nonzero when S_M depends on the linear connection $\Gamma_{\mu\kappa}{}^{\lambda}$. Consider the following choices for matter and radiation Lagrangians: Klein-Gordon, Dirac, Proca, Maxwell, Yang-Mills, and perfect-fluid fields, plus couplings.[27] None of these depends on the $\Gamma_{\mu\kappa}{}^{\lambda}$; therefore, $J^{\mu\kappa}{}_{\lambda}=0$ in all these cases.[28] Since these, or some minor variant, are the normal choices in physical theories of matter, I shall take $J^{\mu\kappa}{}_{\lambda}=0$ in all the following.

It is instructive to compare this set with the set derived from (11), the Einstein Lagrangian. Einstein's field equations are

$$\delta\Gamma_{\mu\kappa}{}^{\lambda}:\ \frac{1}{\kappa}\nabla_{\lambda}g^{1/2}g^{\mu\kappa}=J^{\mu\kappa}{}_{\lambda}=0\,, \tag{36}$$

$$\delta g^{\mu\kappa}:\ \frac{1}{\kappa}g^{1/2}(R_{\mu\kappa}-\tfrac{1}{2}g_{\mu\kappa}R)=g^{1/2}T_{\mu\kappa}\,. \tag{37}$$

Equations (36) imply the Riemannian condition $\nabla_{\mu}g_{\mu\lambda}=0$, so that the connection $\Gamma_{\mu\kappa}{}^{\lambda}$ is just the Christoffel symbol. Equations (37) are Einstein's equations. Note that it is not Yang's equations (15) but rather the second set (25) that correlates with Einstein's equations. Yang's equations correlate with the metric condition (34).

In Einstein's theory, the stress-energy tensor is conserved:

$$\nabla_{\mu}T^{\mu\kappa}_{\cdot\cdot}=0\,. \tag{38}$$

This is a direct result of the Bianchi identity (16) and the metric condition (36). Is $T^{\mu\kappa}_{\cdot\cdot}$ conserved for the Yang theory? To check this, one must introduce the Riemannian covariant derivative D_{μ} constructed using the Christoffel symbols as its connection coefficients which one forms from $g_{\mu\kappa}$. When the geometry is Riemannian, ∇_{μ} and D_{μ} correspond. Because of (38) and when the geometry is Riemannian, the quantity

$$Q\equiv\int_{\Sigma}d\Sigma^{\mu}T_{\mu\kappa}\xi^{\kappa} \tag{39}$$

is independent of the three-hypersurface Σ on which it is defined if the vector field ξ^{κ} satisfies the Killing equation

$$D_{(\mu}\xi_{\kappa)}=0\,. \tag{40}$$

Each independent vector field ξ^{κ} solving (40) is called a Killing vector field. Each ξ^{κ} is the flow of an isometry of $g_{\mu\kappa}$. When the geometry is not Riemannian, Killing's equation (40) is still constructed using the Riemannian covariant derivative D_{μ}. Therefore, in the non-Riemannian geometry of Yang's theory, one must have

$$D_{\mu}T^{\mu\kappa}_{\cdot\cdot}=0 \tag{41}$$

for conserved quantities, as in (39), to exist.

One may compute $D_{\mu}T^{\mu\kappa}$ by expressing D_{μ} in terms of ∇_{μ}. The connections relate via

$$\Gamma_{\mu\lambda}{}^{\kappa}=\{{}_{\mu\lambda}{}^{\kappa}\}+S_{\mu\lambda}{}^{\kappa}+2S^{\kappa}_{\cdot(\mu\lambda)}+Q_{(\mu\lambda)}{}^{\kappa}_{\cdot}-\tfrac{1}{2}Q^{\kappa}_{\cdot\mu\lambda}\,, \tag{42}$$

where $\{{}_{\mu\lambda}{}^{\kappa}\}$ is the Christoffel symbol formed from $g_{\mu\kappa}$:

$$\{{}_{\mu\lambda}{}^{\kappa}\}\equiv\tfrac{1}{2}g^{\kappa\epsilon}(\partial_{\mu}g_{\lambda\epsilon}+\partial_{\lambda}g_{\mu\epsilon}-\partial_{\epsilon}g_{\mu\lambda})\,. \tag{43}$$

Expressing D_{μ} in terms of ∇_{μ} through (42), using the field equations (36), (37), the Bianchi identity in the presence of torsion (21), and the definition of $Q_{\mu\kappa\lambda}$ (13) gives[29]

$$\begin{aligned}\kappa D_{\mu}T^{\mu\lambda}_{\cdot\cdot}&=D_{\mu}H^{\mu\lambda}_{\cdot\cdot}\\&=S_{\kappa\mu}{}^{\tau}R^{\mu\kappa}_{\cdot\cdot\rho}{}^{\sigma}R^{\lambda}_{\cdot\tau\sigma}{}^{\rho}+S^{\lambda}_{\cdot\rho}{}^{\rho}R_{\kappa\mu\rho}{}^{\sigma}R^{\kappa\mu}_{\cdot\cdot\sigma}{}^{\rho}\,.\end{aligned} \tag{44}$$

Thus generally one only gets conservation of stress-energy-momentum if there is no torsion. Because of this, I conclude that we should constrain torsion to be zero. If this is done in the variation, Eq. (19) replaces (34) when $J^{\mu\kappa}{}_{\lambda}=0$. We are forced to break the gauge analogy noted in Sec. I.

We can get rid of the torsion, and apparently doing so is necessary, but can we recoup Riemannian geometry with $Q_{\mu\kappa\lambda}=0$? This is not feasible. If we require Riemannian geometry, Loos's previously noted argument[10,11] on (17) gives Einstein spaces almost always. But $R_{\mu\kappa}=\omega g_{\mu\kappa}$ implies $H^{\mu\kappa}_{\cdot\cdot}=0$, which implies $T^{\mu\kappa}_{\cdot\cdot}=0$. *Therefore, there are almost no matter solutions to Yang's theory which are Riemannian.* Of course, "matter" here means any nongravitational source field. Since $Q_{\mu\kappa\lambda}\neq 0$, Yang's theory falls prey to the difficulties of Weyl's theory.[3,30] Fundamental lengths change under parallel displacement and would thereby vary depending on their history. Elementary particle-mass values are such fundamental lengths. Since particle masses show no such historical effect, the values that $Q_{\mu\kappa\lambda}$ could realistically take on are limited significantly. But any restriction on the magnitude of $Q_{\mu\kappa\lambda}$ similarly restricts the magnitude of $T_{\mu\kappa}$. Conversely, the sharply equal mass values of equivalent particles could be used to place a bound on $Q_{\mu\kappa\lambda}$.

The homogeneous isotropic perfect-fluid metric of Robertson and Walker is used to pose the standard general-relativistic cosmology.[31] If one adds the secondary constraint of Riemannian geometry, Eq. (17) implies that the curvature scalar R is constant. *This is sufficient to force the Robertson-Walker solution to be static.* One rules this out by the observed Hubble red-shift. To get dynamical solutions, one must permit $Q_{\mu\kappa\lambda}\neq 0$. Note also that the Hawking singularity theorems no longer follow from the standard energy assumptions.[32]

One other unusual property of the Stephenson-

Yang theory is that $H^{\mu\kappa}_{\cdot\cdot}$ is traceless. This forces the stress-energy tensor $T^{\mu\kappa}_{\cdot\cdot}$ to be traceless also. Normally the trace of the $T^{\mu\kappa}_{\cdot\cdot}$ for some matter field is proportional to the Lagrangian mass term of the field; in fact, any terms of the Lagrangian with dimensional coupling constants will show up. Unless one wants only massless matter, some prescription must be given to correct this. One could consider introducing conformally invariant masses.[33,34] Replace any mass m by $\bar{m}g^{1/8}$ for constant $\bar{m}$. This changes the coupling to a new dimensionless constant $\bar{m}$, and causes $T^{\mu\kappa}_{\cdot\cdot}$ to be traceless. The full effect of this prescription in field theory has not been worked out.

IV. COMMENTS AND CONCLUSIONS

I admit my intention in looking at the Stephenson-Yang theory was to show that it cannot be a viable alternative to Einstein's theory. But my first conclusions were promising. I ruled out all the unphysical solutions previously considered. Matter-free Riemannian solutions differ from Einstein's only by an insignificant and arbitrary cosmological constant. The theory, though, still has serious difficulties which show up when matter is introduced.

The presence of matter forces us out of a Riemannian geometry. A whole new, additional set of forces in the geodesic equation will arise from $Q_{\mu\kappa\lambda} \equiv -\nabla_\mu g_{\kappa\lambda}$. It is not clear what their effect will be. Posing the matter problem is also beset with the problem of needing $T_\mu{}^\mu_{\cdot} = 0$. For a perfect fluid this is equivalent to relating the density and pressure via $p = \frac{1}{3}\rho$, the normal prescription for a radiation fluid. If one adopts a conformally invariant approach to matter, then our normal interpretation of the stress tensor $T_{\mu\kappa}$ must change. The question of how is not yet answered. Until it is answered, one cannot pose the equivalent of the Tolman-Oppenheimer-Volkov equations of Einstein's theory describing a static, spherically symmetric fluid.[35]

There is one problem that I did not previously mention, which I have saved for last; that is, Yang's theory satisfies Newtonian correspondence only in part. Newtonian correspondence is usually spoken of in two forms. In the first form, one has the Schwarzschild solution which effects geodesics for weak fields in the same manner as Newton's theory. It is this correspondence that implies that the constant G in (24) is the Newtonian gravitation constant. Yang's theory satisfies this. In the second form of Newtonian correspondence, one shows that in the limit of weak static gravitational fields, Einstein's equations reduce to Poisson's equation

$$\nabla^2\phi = 4\pi G\rho. \qquad (45)$$

The ρ is the mass density. The ϕ is identified with the h_{00} component of the weak-field metric $g_{\mu\kappa} = \eta_{\mu\kappa} + h_{\mu\kappa}$ via $h_{00} = -2\phi$. This limit does not work for the Yang theory. Equations (34) and (35) do not reduce to (45). We cannot make the identification of the coupling κ in (10) in terms of G as in Einstein's theory. The weak Yang equations are not even linear because $H_{\mu\kappa}$ is quadratic in $R_{\mu\kappa\lambda}{}^\epsilon$. Terms quadratic in h_{00} and $Q_{\mu\kappa}{}^\lambda_{\cdot}$ are the lowest order. This is a serious difficulty.

The Yang equations with matter will have to be developed further before any other concrete conclusions can be made. But considering all of these difficulties there is one question which needs to be discussed: What good reason does one have to pose a "gauge"-type theory of gravitation when one has a perfectly good theory in Einstein's? I can find only two justifications for a gauge formulation of gravity.

First, there is the hope for a unified theory of weak, electromagnetic, strong, and gravitational forces. Putting the Lagrangian for gravity in quadratic form as in (10) makes gravity more obviously like attempts at unifying the first three.[36]

Second, Yang derives this choice from an "integral formulation" which he feels "is conceptually superior to the differential formalism and allows for natural development of additional concepts."[1] This integral approach "further allows a mathematical and physical discussion of the gravitational field, *as a gauge field*,"[1] resulting in the equations I have discussed.

Yang's integral formulation is *not* "conceptually superior"[1] to the standard formulation in terms of differential geometric concepts. It is rather only a rewrite of results related to the study of the holonomy group in the standard differential formulation.[37] The two approaches are equivalent; his formalism no more strongly leads to the quadratic gravitational Lagrangian (10) than the traditional approach. Utiyama first noted the similarity between gauge-field concepts and the concepts of Riemannian geometry in Einstein's theory, and called Einstein's theory a gauge theory.[17] This is because the kinematics of gauge theories and general relativity are the same. This has been long known by those who study the two from the unifying point of view of fiber bundles.[17-19] Yang believes it "an unnatural interpretation of gauge fields" to call Einstein's theory a gauge theory.[38] He reserves the term for use only by those theories having a dynamics following from the Lagrangian quadratic in the gauge curvature, given by (6). For an arbitrary gauge space erected on space-time to house the gauge-tensor fields, Uzes's theorem[14,15] indicates that the quadratic choice is best. But if the gauge space should have additional structure, then other interesting choices for dynamical statements are available. The tangent space, whose gauge field gives rise to gravity, has such additional structure; it is "soldered" to the mainfold, to use a term of Trautman,[19] in a way no other gauge space is. This permits us, only in this case, to have a Lagrangian linear in the gauge curvature and yet to have proper hyperbolic field equations. An additional example of a special case is the Dirac spin space. The linkage of the spin space to the tangent space via the Dirac matrices lets one have standard gauge fields, but with *masses* without introducing any symmetry breakdown.[34,39] It is because these are special cases that one would probably *not* want to take the "general" prescription.

Also, the different character of gravity is felt by the way the metric interacts with matter in a characteristically nonpolynomial way.[40] A factor of $g^{1/2}$ coming from the invariant measure multiplies each term of the Lagrangian. Interaction via the connection $\Gamma_{\mu\kappa}{}^{\lambda}$ in a minimal coupling through the covariant derivative is characteristically not important. Minimal coupling is normally the only type of interaction for a gauge field. But it is just this difference which to me sheds doubt on a single unified theory of weak, electromagnetic, strong, and gravitational forces in a theory purely of the Weinberg-Salam type.[35]

Independent of these somewhat philosophical considerations, one has a number of difficulties to be explicitly resolved to make a viable theory: (1) How can the full Yang equations be made to reduce to Newton's? (2) What is the strength of the coupling constant κ in this theory? (3) Can we live with a traceless stress-energy tensor? (4) How does one treat the necessity of having $\nabla_\mu g_{\kappa\lambda} \neq 0$ when matter is present? What interpretation are these added degrees of freedom to have, since they show up in the geodesic equation of a test particle? (5) How should homogeneous, isotropic cosmological solutions act? (6) Although Yang's theory satisfies the three traditional tests, does it satisfy the full range of gravitational tests? (7) Since the interaction with matter is so different in Yang's and Einstein's theories, one should be able to construct some experiment which will differentiate the two. What would one be? These are only a few of the questions that must be posed and answered definitively in favor of the theory to settle firmly the question of the viability of this type of gravitational theory.

ACKNOWLEDGMENTS

I would like to thank Richard P. Treat for discussions which led to this research and for pointing out the relevance of Ref. 11 to the Yang theory.

[1]C. N. Yang, Phys. Rev. Lett. 33, 445 (1974).
[2]G. Stephenson, Nuovo Cimento 9, 263 (1958).
[3]Hermann Weyl, *Space—Time—Matter* (Dover, New York, 1952), p. 309.
[4]A. H. Thompson, Phys. Rev. Lett. 34, 507 (1975); 35, 320 (1975).
[5]Richard Pavelle, Phys. Rev. Lett. 34, 1114 (1975).
[6]A. Palatini, R. C. Circ. Mat. Palermo 43, 203 (1919), referenced in W. Pauli, *Theory of Relativity* (Pergamon, New York, 1958), pp. 214–15.
[7]J. A. Schouten, *Ricci-Calculus* (Springer, Berlin, 1954).
[8]J. A. Schouten (see Ref. 7, Chap. III) gives a full terminology to deal with non-Riemannian geometries. A manifold with metric $g_{\mu\kappa}$ has a connection $\Gamma_{\mu\kappa}{}^{\lambda}$ which is *symmetric*, if the torsion $S_{\mu\kappa}{}^{\lambda} \equiv \Gamma_{[\mu\kappa]}{}^{\lambda}$ is zero; otherwise it is *nonsymmetric*. The connection is *metric* if $Q_{\mu\kappa\lambda} \equiv -\nabla_\mu g_{\kappa\lambda}$ is zero, or *semimetric* if $Q_{\mu\kappa\lambda} = A_\mu g_{\kappa\lambda}$; otherwise it is *nonmetric*. A Riemannian geometry has a metric-symmetric connection. A Riemann-Cartan geometry has a metric-nonsymmetric connection. A Weyl geometry has a semimetric-symmetric connection. A Weyl-Cartan geometry has a semimetric-nonsymmetric connection. For Yang's theory one must have a nonmetric connection, but it may be either symmetric or nonsymmetric. The field equations do not constrain the connection to be metric as with Einstein's theory.
[9]C. W. Kilmister and D. J. Newman, Proc. Camb. Philos. Soc. 57, 851 (1961); C. W. Kilmister, *Les Theories Relativistes de la Gravitation* (Centre National de la Recherche Scientifique, Paris, 1962).
[10]H. G. Loos, Ann. Phys. (N.Y.) 25, 91 (1963).
[11]H. G. Loos and R. P. Treat, Phys. Lett. 26A, 91 (1967).
[12]At the end of his paper (see Ref. 1), Yang discusses a variational principle, varying the connection and metric as in my specification. His variation, though, is rather some construct to show that a Riemannian choice for the geometry can be justified. A variation such as this does not generalize to the source case, and I find it difficult to see the point of it.
[13]C. N. Yang and R. L. Mills, Phys. Rev. 96, 191 (1954).
[14]For notation, I will follow most closely Hendricus G. Loos, J. Math. Phys. 8, 2114 (1967).
[15]In fact, $\phi_{\mu\kappa i}{}^{j}$ is a preeminent isotensor concomitant of $\Phi_{\mu i}{}^{j}$ because all isotensor concomitants of $\Phi_{\mu i}{}^{j}$ and its partial derivatives are algebraic concomitants of $\phi_{\mu\kappa i}{}^{j}$ and its covariant derivatives.
[16]C. A. Uzes, Ann. Phys. (N.Y.) 50, 534 (1968), has shown that all Lagrangians constructed for $\Phi_{\mu i}{}^{j}$ give rise to trivial gauge charges except for (6).
[17]Ryogu Utiyama, Phys. Rev. 101, 1597 (1956).
[18]T. W. B. Kibble, J. Math. Phys. 2, 212 (1961).
[19]A. Trautman, Rep. Math. Phys. 1, 29 (1970).
[20]The set of all possible coordinate changes on space-time is called its diffeomorphism group $\mathfrak{D}$. One connects elements of $\mathfrak{D}$ with elements of gauge $GL(4,R)$

as noted in the text. This is a homomorphism of $\mathfrak{D}\rightarrow$ gauge GL(4, R) which is neither one-to-one nor onto. The groups are structurally very different, and I refer the reader to DeWitt for details; see Bryce S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965). All indices on the connection $\Gamma_{\mu\kappa}{}^{\lambda}$ and curvature $R_{\mu\kappa\lambda}{}^{\epsilon}$ are now in the tangent bundle; whereas, in the general case, these indices are split between the gauge bundle and the tangent bundle. Therefore, more scalar invariants exist, implying more choices for Lagrangian dynamical statements. The third subtlety is the preeminence of the tangent space metric $g_{\mu\kappa}$, over the metric in any other bundle, in forming the invariant measure in space-time, as I noted in the text.

[21]Since one also has no proof of the Birkhoff theorem (see Ref. 25, p. 337), there is no requirement that one must choose the exterior solution of the gravitation field of the sun to be static, nor does one know if it must be unique.

[22]This identity can be proved using 2-spinor techniques and is given by Pirani as an exercise; see A. Trautman, F. A. E. Pirani, and H. Bondi, *Lectures on General Relativity* (Prentice-Hall, New Jersey, 1965), p. 317. I know of no direct proof using only the symmetries and identities of $C_{\mu\kappa\lambda}{}^{\epsilon}$.

[23]Although I have only proved this conjecture in four dimensions for a metric of signature -2, I conjecture that it is true for other dimensions and signatures.

[24]See Ref. 7, Sec. II.9.

[25]Steven Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), p. 146.

[26]For example, see Ref. 25, pp. 179–182.

[27]One poses these usually Minkowski space systems simply by letting the partial derivatives go over to covariant derivatives and using the invariant measure to construct the action.

[28]In the Klein-Gordon case the covariant derivative of the scalar field equals the partial derivative of it. For the Proca and Maxwell fields it enters as $\nabla_{[\mu}A_{\kappa]} = \partial_{[\mu}A_{\kappa]}$. For the Dirac case, $\nabla_{\mu}g_{\kappa\lambda}\neq 0$ implies $\nabla_{\mu}\gamma_{\kappa}\neq 0$ (γ_{μ} are the Dirac matrices) which decouples the usual dependence of the spin connection on $\Gamma_{\mu\kappa}{}^{\lambda}$. The Yang-Mills and perfect-fluid Lagrangians simply do not depend on $\Gamma_{\mu\kappa}{}^{\lambda}$. The conformal Klein-Gordon equation, which contains an additional $\frac{1}{12}R\phi^2$ term in the Lagrangian and gives $J^{\mu\kappa}{}_{\lambda}\neq 0$, has been excluded since it must also be excluded in the Einstein case when a Palatini variation is used.

[29]That this equation is not zero when $T_{\mu\kappa}=0$ indicates that $H_{\mu\kappa}=0$ has nontrivial integrability conditions which will restrict the torsion when it is present.

[30]A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge Univ. Press, Cambridge, 1965).

[31]See Ref. 25, Chaps. 14 and 15.

[32]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge Univ. Press, London, 1973).

[33]T. Fulton, F. Rohrlich, and L. Witten, Rev. Mod. Phys. 34, 442 (1962).

[34]Edward E. Fairchild, Jr., Ph.D. thesis, University of Texas at Austin, 1975 (unpublished).

[35]See Ref. 25, Chap. 11.

[36]Steven Weinberg, Rev. Mod. Phys. 46, 225 (1974), and references therein.

[37]See Ref. 13 and Ref. 7, Secs. VII. 4–6.

[38]See footnote 5 of Ref. 1.

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

[40]Abdus Salam, in *Nonpolynomial Lagrangians, Renormalisation and Gravity*, Vol. 1 of proceedings of the 1971 Coral Gables Conference on Fundamental Interactions at High Energy, edited by M. Dal Cin, G. J. Iverson, and A. Perlmutter (Gordon and Breach, New York, 1971), p. 3.