Pure Connection Action Principle for General Relativity

Kirill Krasnov
School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
(Received 23 March 2011; published 24 June 2011)

It has already been known for two decades that general relativity can be reformulated as a certain gauge theory, so that the only dynamical field is an SO(3) connection and the spacetime metric appears as a derived object. However, no simple action principle realizing these ideas has been available. A new elegant action principle for such a "pure connection" formulation of GR is described.

DOI: 10.1103/PhysRevLett.106.251103 PACS numbers: 04.20.Cv

In [1] Plebanski has shown that instead of the spacetime metric the dynamical field of general relativity (GR) can be taken to be a collection (triple) of two-forms satisfying a certain algebraic equation. This idea was taken further in [2,3], where it was shown that the two-form field (as well as the Lagrange multiplier field of the Plebanski formulation) can be integrated out to obtain a "pure connection formulation" of GR. An action principle realizing this idea in the case $\Lambda = 0$ of zero cosmological constant GR was described in [2]. The case $\Lambda \neq 0$ proved to be more difficult, and only a rather complicated (and erroneous, see [4]) action was given in [3]. The correct treatment of the $\Lambda \neq 0$ case was given in [5] and then [6]. The resulting actions, however, are far from simple, which raises serious doubts about the usefulness of reformulations of this type. The purpose of this Letter is to point out that an elegant and simple "pure connection" action principle for GR encompassing both $\Lambda \neq 0$ and $\Lambda = 0$ cases is possible. Our action principle also provides an answer to the question posed in [5] as to "what it is that makes the Lagrangians found better than their obvious generalizations." Indeed, we shall see that the GR Lagrangian is the unique Lagrangian (within a certain class) possessing a property that its SO(3) gauge field is the self-dual part of the metric-compatible spin connection. Another significant improvement achieved by our work is that in contrast to [2,5,6], our action is a functional of the connection only; no additional auxiliary field is necessary.

We start by describing the new variational principle, and then prove that solution of the arising Euler-Lagrange equations are in one-to-one correspondence with solutions of Einstein's theory. We first discuss the case of the Riemannian signature GR, of importance, in particular, in the branch of mathematics studying Einstein manifolds, see, e.g., [7]. It is simpler, for the connections that one deals with are real-valued. A pure connection formulation is also possible for the physically relevant Lorentzian signature GR, but, similar to the Plebanski formulation [1], it uses complex-valued $SO(3,\mathbb{C})$ connections and requires certain reality conditions to be added. We discuss this at the end of the Letter.

As in [2,3], the main dynamical field of our formulation is an SO(3) (real-valued for Riemannian signature case) connection field A^i, $i = 1, 2, 3$ over the spacetime manifold M. In this Letter we only discuss the local aspects, so we leave the choice of an SO(3) bundle over M unspecified. We just mention that in case M is compact a very specific bundle must be chosen in order to obtain the equivalence to GR. We take A^i to be a dimensionful quantity with dimensions of $A \sim 1/L$, where L is a unit of length. Given a connection A^i, its curvature is given by $F^i = dA^i + (1/2)\epsilon^{ijk} A^j \wedge A^k$. Here the form notation is used, and $\wedge$ denotes the wedge product of forms. As in [2,3], we consider the 4-form $F^i \wedge F^j$, which is valued in the second (symmetric) power of the Lie algebra $\mathfrak{su}(2)$. Using the density weight one antisymmetric tensor $\tilde{\epsilon}^{\mu\nu\rho\sigma}$, which does not need a metric for its definition (here $\mu, \nu, \ldots$ are the spacetime indices), we can convert the 4-form $F^i \wedge F^j$ into a density weight one symmetric 3×3 matrix $\tilde{X}^{ij} := (1/4)\tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} F^j_{\rho\sigma}$, so that $F^i \wedge F^j = \tilde{X}^{ij} d^4x$. We note that $\tilde{X}^{ij}$ has dimensions of $1/L^4$. Now consider an arbitrary homogeneous of degree one, gauge invariant function f:Mat$(3 \times 3) \to \mathbb{R}$, i.e., a function satisfying $f(\alpha\tilde{X}) = \alpha f(\tilde{X})$ as well as $f(O\tilde{X}O^T) = f(\tilde{X})$, $O \in SO(3, R)$. Then $f(\tilde{X})$ is a density weight one, and can be integrated over the spacetime to produce an action. We refer the reader to, e.g., [8] for more details on this construction of diffeomorphism invariant actions. We also note that the sketched construction of actions is somewhat similar to that described in [9] in the context of stable differential forms.

The simplest possible diffeomorphism invariant gauge theory action corresponds to $f(\tilde{X}) = \mathrm{Tr}(\tilde{X})$. This, however, gives a topological theory without any interesting dynamics. As we shall now see, general relativity (with $\Lambda \neq 0$) arises for a certain other choice of f. To describe it, let us somewhat restrict the class of connections that are used in the variational principle. Thus, following [10] we call a connection A^i *definite* if the corresponding real symmetric matrix $\tilde{X}^{ij}$ is definite, i.e., its eigenvalues are all nonzero and of the same sign (everywhere in M). For such connections we have a well-defined notion of the matrix square root of $\tilde{X}^{ij}$, which is a symmetric matrix $(\sqrt{\tilde{X}})^{ij}$ such that $(\sqrt{\tilde{X}})^{ij}(\sqrt{\tilde{X}})^{jk} = \tilde{X}^{ik}$. Explicitly, if the matrix $\tilde{X}^{ij}$

is diagonalized by an orthogonal transformation $O \in \mathrm{SO}(3)$, i.e., $\tilde{X} = ODO^T$, $D = \mathrm{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3)$, then $(\sqrt{\tilde{X}}) = O\sqrt{D}O^T$, $\sqrt{D} = \mathrm{diag}(\sqrt{\tilde{\lambda}_1}, \sqrt{\tilde{\lambda}_2}, \sqrt{\tilde{\lambda}_3})$. This involves a choice of the branch of the square root function. For our purposes any of the two branches can be taken; the action is independent of this choice. Indeed, consider the function

$$f(\tilde{X}) := \frac{1}{16\pi G\Lambda}(\mathrm{Tr}\sqrt{\tilde{X}})^2. \tag{1}$$

Here G, Λ are the Newton's and cosmological constants, respectively. Note that because of the second power present here the function (1) is independent of which branch of the square root is used. The function (1) is homogeneous of degree one and gauge invariant. Thus, it satisfies all the requirements discussed above, and so $f(\tilde{X})$ can be integrated over the spacetime to produce an action. We note that in the units $c = 1$ used in this Letter the quantity $1/(G\Lambda)$ has dimensions of $\hbar$. Thus, when (1) is integrated over the manifold the result will have dimensions of $\hbar$, as is appropriate for the action.

Having in mind the construction just described, and switching to the compact form notation, we can write our action functional as follows:

$$S_{\mathrm{GR}}[A] = \frac{1}{16\pi G\Lambda}\int_M (\mathrm{Tr}\sqrt{F^i \wedge F^j})^2. \tag{2}$$

This is the action principle that is the subject of this Letter.

To prove that (2) is indeed an action for general relativity in disguise, we need to find the corresponding Euler-Lagrange equations. The matrix of partial derivatives of the function (1) with respect to the components of the matrix $\tilde{X}^{ij}$ is given by

$$\frac{\partial f}{\partial \tilde{X}^{ij}} = \frac{1}{16\pi G\Lambda}(\mathrm{Tr}\sqrt{\tilde{X}})(\sqrt{\tilde{X}^{-1}})^{ij}. \tag{3}$$

Note that the inverse of $\tilde{X}^{ij}$ exists as is guaranteed by the restriction that definite connections are considered. Note also that, as is appropriate for a function of degree of homogeneity one, we have

$$\frac{\partial f}{\partial \tilde{X}^{ij}}\tilde{X}^{ij} = f(X). \tag{4}$$

If we now define

$$B^i := \frac{\partial f}{\partial \tilde{X}^{ij}}F^j, \tag{5}$$

then the Euler-Lagrange equations for (2) read

$$D_A B^i = 0. \tag{6}$$

Here, D_A is the covariant derivative with respect to the connection A^i. We note that (6) is a set of 3×4 second-order differential equations for the 3×4 components of the connection A^i.

We now show that (6), together with the definition (5) of the two-form field B^i are equivalent to the field equations of Plebanski formulation of GR [1]. To this end we first note that the two-form field (5) satisfies a set of algebraic equations. Indeed, using the definition of $\tilde{X}^{ij}$ we have

$$\frac{1}{4}\tilde{\epsilon}^{\mu\nu\rho\sigma}B^i_{\mu\nu}B^j_{\rho\sigma} = \frac{\partial f}{\partial \tilde{X}^{ik}}\frac{\partial f}{\partial \tilde{X}^{jl}}\tilde{X}^{kl} = \left(\frac{\mathrm{Tr}\sqrt{\tilde{X}}}{16\pi G\Lambda}\right)^2\delta^{ij}.$$

Thus, the two-form field (5) satisfies

$$B^i \wedge B^j \sim \delta^{ij}, \tag{7}$$

which is the basic equation of Plebanski's formulation of GR [1]. The Einstein equations then arise as follows. Given a triple of two forms B^i satisfying (7) there is a canonically defined real Riemannian signature spacetime metric (determined by the condition that B^i (or F^i) are self-dual, and that a multiple of $B^i \wedge B^i$ is the volume form). An explicit formula for this metric can be given as an expression cubic in B^i, but will not be needed here. The Eqs. (6) can be solved for the connection components A^i and the Eq. (7) implies that the resulting SO(3) connection is the self-dual part of the metric-compatible spin connection. The Eq. (5) rewritten as

$$F^i = \left(\frac{\partial f}{\partial \tilde{X}^{ij}}\right)^{-1}B^j \tag{8}$$

then implies that the curvature of the self-dual part of the spin connection is self-dual as a two-form, which is equivalent to the Einstein condition $R_{\mu\nu} \sim g_{\mu\nu}$. For more details on Plebanski formulation in the notations close to those of this Letter the reader is referred to [11]. We have thus shown that the (definite) solutions of our theory are in one-to-one correspondence with solutions of general relativity. Indeed, the above discussion shows that any (definite) connection satisfying the field Eqs. (6) gives rise to an Einstein metric (obtained by requiring F^i to be self-dual). In the other direction, any solution of Einstein's theory gives rise to a (definite) solution of the theory (2) (by considering the self-dual part of the metric-compatible spin connection).

We note that the logic of the above proof of equivalence to GR could be reversed, and one could *derive* (1) as the only function of the matrix $\tilde{X}^{ij}$ of curvature wedge products such that the corresponding action produces Plebanski field equations. Indeed, it is clear that the key point about the particular choice (1) is that it leads to (7). This is the case for

$$\frac{\partial f}{\partial \tilde{X}^{ij}} \sim (\mathrm{Tr}\sqrt{\tilde{X}})(\sqrt{\tilde{X}^{-1}})^{ij}, \tag{9}$$

where we need the trace prefactor in order to guarantee that $\partial f/\partial\tilde{X}$ is of degree of homogeneity zero. This then integrates to (1).

So far we have discussed the case of GR with nonzero cosmological constant. Indeed, in the limit $\Lambda \to 0$ the action (2) is singular. However, in this limit the (exponential of the) action present in the quantum mechanical path integral of the theory can be viewed as a delta function imposing the constraint

$$\mathrm{Tr}\sqrt{F^i \wedge F^j} = 0. \tag{10}$$

This is the same equation as was found in [2] by rewriting the $\Lambda = 0$ general relativity in the pure connection formalism. Indeed, the condition that the trace of the square root of a matrix is equal to zero can be rewritten as an equation on the matrix itself. For this we, as in [3], denote $Y^{ij} \sim \sqrt{F^i \wedge F^j}$, and write down the characteristic equation for Y:

$$Y^3 - \mathrm{Tr}(Y)Y^2 + \tfrac{1}{2}([\mathrm{Tr}(Y)]^2 - \mathrm{Tr}(Y^2))Y = \det(Y).$$

Now, multiplying by Y, taking the trace and using $\mathrm{Tr}(Y) = 0$ we get

$$\mathrm{Tr}\,(Y^4) - \tfrac{1}{2}[\mathrm{Tr}(Y^2)]^2 = 0. \tag{11}$$

Rewriting this in terms of $X = Y^2$, $X^{ij} \sim F^i \wedge F^j$ we get

$$\mathrm{Tr}\,(X^2) - \tfrac{1}{2}[\mathrm{Tr}(X)]^2 = 0, \tag{12}$$

which is just the condition found in [2,3]. Thus, in the sense described, our action (2) encompasses both $\Lambda \neq 0$ and $\Lambda = 0$ cases.

Finally, we present an alternative derivation of the action (2) directly from the Plebanski formulation of GR, via the same procedure as is followed in [3,5,6]. However, unlike in these references we use a different set of invariants of 3×3 matrices to solve for (eigenvalues), which allows us to obtain a much more compact final result. In particular, we are able to integrate over all the fields apart from the connection, while the actions in [5,6], still contain an auxiliary field. The Plebanski action for (Riemannian signature) GR with a cosmological constant Λ is a functional of the connection A^i, an $\mathfrak{su}(2)$-valued two-form B^i, as well as a field of Lagrange multipliers Ψ^{ij}. It is given by

$$S_{\mathrm{Pleb}} = \frac{1}{8\pi G}\int\left[B^i \wedge F^i - \frac{1}{2}\left(\Psi^{ij} + \frac{\Lambda}{3}\delta^{ij}\right)B^i \wedge B^j\right].$$

More details on this formulation can be found in, e.g., [11]. Integrating out the two-form field one gets the following action

$$S[A, \Psi] = \frac{1}{16\pi G}\int\left(\Psi^{ij} + \frac{\Lambda}{3}\delta^{ij}\right)^{-1} F^i \wedge F^j, \tag{13}$$

where, as in [5,6], it is assumed that the matrix $(\Psi^{ij} + (\Lambda/3)\delta^{ij})$ is invertible. It is now convenient to rescale the Lagrange multipliers field and write the action as

$$S[A, \tilde{\Psi}] = \int(\tilde{\Psi}^{ij} + \alpha\delta^{ij})^{-1} F^i \wedge F^j, \tag{14}$$

where

$$\alpha := \frac{16\pi G\Lambda}{3}, \tag{15}$$

in units $\hbar = c = 1$, is a dimensionless quantity. Note that $\alpha \sim M_\Lambda^2/M_p^2$ and so, for the observed value of the cosmological constant, is of the order $\alpha \sim 10^{-120}$.

In the final step we integrate out the Lagrange multiplier field $\tilde{\Psi}^{ij}$. Let us drop the tilde on the symbol for brevity. We can the rewrite the above action as

$$S[A, \Psi] = \int(\mathrm{vol})\mathrm{Tr}((\Psi + \alpha Id)^{-1}X), \tag{16}$$

where we have introduced $F^i \wedge F^j = (\mathrm{vol})X^{ij}$, and (vol) is an arbitrary auxiliary four form on our manifold. To integrate out the matrix Ψ we have to solve the field equations for it, and then substitute the result back into the action. Assuming that the solution for Ψ can be written as a matrix function of X, we conclude that Ψ will be diagonal if X is. Thus, we can simplify the problem of finding Ψ by using an SO(3) rotation to go to a basis in which X is diagonal. We then look for a solution in which Ψ is also diagonal. Denoting by $\lambda_1, \lambda_2, \lambda_3$ the eigenvalues of X^{ij}, and by a, b, $-(a+b)$ the components of the diagonal matrix Ψ, we get the following action functional to consider

$$F[a, b, \lambda] = \frac{\lambda_1}{\alpha + a} + \frac{\lambda_2}{\alpha + b} + \frac{\lambda_3}{\alpha - (a+b)}. \tag{17}$$

We now have to vary this with respect to a, b and substitute the solution back to obtain the defining function as a function of λ_i. Assuming that no of the denominators in (17) are zero we get the following two equations

$$\begin{aligned}(\alpha + a)^2\lambda_3 &= (\alpha - (a+b))^2\lambda_1,\\ (\alpha + b)^2\lambda_3 &= (\alpha - (a+b))^2\lambda_2.\end{aligned} \tag{18}$$

Taking the (positive branch of the) square root and adding the results we get $a + b$, which is most conveniently written as

$$\alpha - (a+b) = 3\alpha\frac{\sqrt{\lambda_3}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}}. \tag{19}$$

The other two combinations that appear in (17) are given by similar expressions. It is now clear that the sought function of the matrix X is given by

$$f_{\mathrm{GR}}(\lambda) = \frac{1}{3\alpha}(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})^2 = \frac{1}{3\alpha}(\mathrm{Tr}\sqrt{X})^2.$$

Integrated over the spacetime manifold this is just our action (2). This concludes our proof of the classical equivalence of General Relativity with a nonzero cosmological constant and the theory of connections (2).

Let us now briefly discuss the modifications necessary to extend the above pure connection description of gravity to the Lorentzian signature setting. In this case the self-dual connections are SO(3, $\mathbb{C}$) valued, and so all field become complex valued. The action principle (2) must then be supplemented with appropriate reality conditions. These are similar to those in [1,3], and read

$$F^i \wedge (F^j)^* = 0, \qquad \mathrm{Re}(F^i \wedge F^i) = 0. \tag{20}$$

The first set here says that the subspace in the space of two-forms spanned by F^i is wedge-orthogonal to the complex conjugate subspace, which then implies that the conformal metric defined by declaring F^i to be self-dual is real Lorentzian. The second condition in (20) says that the four-form $F^i \wedge F^i$ is purely imaginary, which guarantees the full metric (including the conformal factor) is real. Note that there are 10 reality conditions in (20), exactly the number that is needed to require 10 metric components to be real.

We conclude with a number of remarks. First, the formulation (2) can be used as the starting point for a new type of the gravitational perturbation theory. Here one expands the action around the constant curvature background, and the usual linearized GR solutions (gravitons) can be seen to appear [8]. It would be very interesting to develop this line of thought further and compute the graviton scattering amplitudes as well as loop corrections using this formalism. Work on these issues is in progress.

Another (potentially important) point about the formulation (2) is that it immediately allows for a very large class of generalizations. Indeed, we have seen that the construction of the action goes through for any homogeneous order one and gauge invariant function $f(\tilde{X})$. The function in (2) is special because it guarantees (7), but other functions can be considered. A large class of diffeomorphism invariant SU(2) gauge theories is then possible, see, e.g., [8,12] for earlier work. A very interesting feature of all these new theories is that they describe just two propagating degrees of freedom, see [13], exactly like general relativity. These theories could be of importance for understanding the ultraviolet behavior of gravity; see [8] for more details on such potential applications.

Note also that it is a very interesting point about (2) that it requires a nonzero Λ. Given that there is now a strong observational evidence for a nonzero cosmological constant, this seems to be a step in the right direction as compared to the usual metric based GR whose action principle works equally well with or without Λ.

Apart from possible applications in quantum gravity, the new formulation (2) may prove instrumental in the classical domain. One promising direction appears to be to questions about the moduli spaces of Einstein metrics on four manifolds, see, e.g., [7], Chap. 12. The point is that the linearization of the new functional (2) behaves differently with respect to diffeomorphisms than the linearization of the Einstein-Hilbert functional. Indeed, one can show that the linearized action is simply independent of certain components of the connection (those which can be changed by an action of a diffeomorphism). This is completely different from the case of the Einstein-Hilbert functional, where diffeomorphisms need to be gauge fixed in a rather nontrivial fashion. Optimistically, this different behavior may make some open rigidity questions about Einstein metrics easier to tackle.

Apart from the above positive features, the new formulation (2) has some difficulties that must be mentioned. The first and foremost is that for applications in physics one needs to know how all other matter couples to gravity. It is not easy to describe this once gravity has been reformulated as a theory of connections. Fortunately, a simple way to couple the usual Yang-Mills gauge fields exists, see [14,15], and also [16] for earlier work. The main idea here is that enlarging the gauge group appropriately and expanding the theory around the constant curvature background in the gravitational sector, one finds the usual Yang-Mills action functional as describing the low energy physics of the nongravitational gauge fields. It is much more difficult to couple to (2) fermionic matter, and it is clear that new ideas will be required here. Work on this is in progress.

Another difficulty with the formulation (2), as well as with the original Plebanski formulation [1], is that the connection field is required to be complex valued (if one is to reproduce the Lorentzian signature sector of GR). Then reality conditions (20) need to be imposed, so that the action (2) is varied among the gauge fields satisfying (20). For the classical theory this is not much of a problem, but if one wants to base on (2) a quantum mechanical treatment, one has to take into account (20) in the path integral, which is a difficult task. One possible way around this problem could be to resort to the analytic continuation to the Riemannian signature metrics, where the reality is trivial to impose. However, it is not at all clear if there is a consistent way to do this in the quantum theory. More work on these issues is required.

The author would like to thank Ingemar Bengtsson for pointing out the reference [4], and an anonymous referee for pointing out works [5,6].

[1] J. F. Plebanski, J. Math. Phys. (N.Y.) **18**, 2511 (1977).
[2] R. Capovilla, T. Jacobson, and J. Dell, Phys. Rev. Lett. **63**, 2325 (1989).
[3] R. Capovilla, T. Jacobson, and J. Dell, Classical Quantum Gravity **8**, 59 (1991).
[4] R. Capovilla, T. Jacobson, and J. Dell, Classical Quantum Gravity **9**, 1839 (1992).
[5] P. Peldan, Classical Quantum Gravity **8**, 1765 (1991).
[6] R. Capovilla and T. Jacobson, Mod. Phys. Lett. A **7**, 1871 (1992).
[7] Arthur L. Besse, *Einstein Manifolds* (Springer-Verlag, Berlin, 2008).
[8] K. Krasnov, arXiv:1101.4788.
[9] N. J. Hitchin, arXiv:math/0107101.
[10] J. Fine (to be published).
[11] K. Krasnov, Gen. Relativ. Gravit. **43**, 1 (2010).
[12] I. Bengtsson, Phys. Lett. B **254**, 55 (1991).
[13] K. Krasnov, Phys. Rev. Lett. **100**, 081102 (2008).
[14] A. Torres-Gomez and K. Krasnov, Phys. Rev. D **81**, 085003 (2010).
[15] A. Torres-Gomez, K. Krasnov, and C. Scarinci, Phys. Rev. D **83**, 025023 (2011).
[16] S. Chakraborty and P. Peldan, Phys. Rev. Lett. **73**, 1195 (1994).