Derivation of the Einstein-Hilbert Action

Abstract

Most people justify the form of the E-H action by saying that it is the simplest scalar possible. But simplicity, one can argue, is a somewhat subjective and ill-defined criterion. Also, simplicity does not shed light on the axiomatic structure of general relativity. Instead, we use the principles of general relativity (i.e. the equivalence principle and the principle of general covariance) plus one more natural asumptiom to derive the E-H action.

1 Introduction

Let us assume that all dynamical properties of our space-time are comprised in the metric tensor $g_{ab}(x)$ which, at the same time, characterizes the behaviour of measuring apparatus. Thus, in a field-like theory, the metric tensor together with its derivatives to finite order can be taken as *dynamical* type variables provided an appropriate *scalar Lagrangian* can be constructed out of them:

$$\mathcal{L}(x) = \mathcal{L}(g_{ab}, \partial_c g_{ab}, \partial_c \partial_d g_{ab}, \cdots). \tag{1.1}$$

We always have in mind the group of general coordinate transformations subject to appropriate differentiability conditions (the manifold mapping group of diffeomorphisms). Since tensors form representations of this group, they are natural objects to serve as the building blocks of a generally covariant physical theory. To ensure the general covariance of the resulting theory, we need to form an invariant *action* integral, so that the statement $\delta S[g] = 0$ is generally covariant, and so also is the dynamics derived from this statement. However, the integral $\int_D d^4x\, \mathcal{L}(x)$, over an invariantly fixed domain D, would not be an invariant as long as $\mathcal{L}(x)$ is a scalar quantity, because

$$\int_D d^4x\, \mathcal{L}(x) = \int_{\bar{D}} d^4\bar{x}\, (\frac{\partial x}{\partial \bar{x}}) \bar{\mathcal{L}}(\bar{x}) \neq \int_{\bar{D}} d^4\bar{x}\, \bar{\mathcal{L}}(\bar{x}), \tag{1.2}$$

when the Jacobian $\frac{\partial x}{\partial \bar{x}} \neq 1$. So we need some quantity $a(x)$ such that

$$\int_D d^4x\, a(x)\, \mathcal{L}(x) = \int_{\bar{D}} d^4\bar{x}\, \bar{a}(\bar{x})\, \bar{\mathcal{L}}(\bar{x}), \tag{1.3}$$

holds, i.e., $a(x)$ needs to be a scalar density transforming as

$$\bar{a}(\bar{x}) = \frac{\partial x}{\partial \bar{x}}\, a(x) \tag{1.4}$$

Since the metric, which transforms as

$$\bar{g}_{ab}(\bar{x}) = \frac{\partial x^c}{\partial \bar{x}^a} \frac{\partial x^d}{\partial \bar{x}^b} \, g_{cd}(x), \tag{1.5}$$

is the only quantity at our disposal, we must construct the scalar density $a(x)$, out of it. By taking the determinants of both sides of the transformation law of the metric tensor, we find

$$\bar{g}(\bar{x}) = \left(\frac{\partial x}{\partial \bar{x}}\right)^2 g(x) \tag{1.6}$$

Since Lorentz signature implies $g = \det(g_{ab}) < 0$, hence

$$\sqrt{-\bar{g}} = \frac{\partial x}{\partial \bar{x}} \sqrt{-g} \tag{1.7}$$

Thus, if we take $a(x) = \sqrt{-g}$ our invariant action becomes,

$$S[g_{ab}] = \int_D d^4x \, \sqrt{-g} \, \mathcal{L}(x). \tag{1.8}$$

By varying this action with respect to the metric tensor, we get the following E-L equation of motion;

$$\frac{\partial \hat{\mathcal{L}}}{\partial g_{ab}} - \partial_c \left(\frac{\partial \hat{\mathcal{L}}}{\partial(\partial_c g_{ab})} \right) + \partial_c \partial_d \left(\frac{\partial \hat{\mathcal{L}}}{\partial(\partial_c \partial_d g_{ab})} \right) - \cdots = 0, \tag{1.9}$$

where $\hat{\mathcal{L}}(x) = \sqrt{-g} \, \mathcal{L}(x)$.

2 Deriving the form of the Lagrangian

This will be based on the following principles;

1. The principle of equivalence: *At every point in an arbitrary curved space-time, we can choose a locally inertial frame in which the laws of physics take the same form as in a global inertial frame of flat space-time*: at any point p, one can choose a coordinate system such that $g_{ab}(p) = \eta_{ab}$ and $\partial_c g_{ab}(p) = 0$. The principle states that in the neighbourhood of this point, the physics is Lorentzian.

2. The assumption: Since almost all of the differential equations of physics are second order, it seems natural to assume that *the metric tensor obeys a second order partial differential equation.*

3. The principle of general covariance: *The form of physical laws is invariant under the group of general coordinate transformations.* The covariance (form invariance) means that the laws of physics must be tensorial. To apply the principle, we need a mathematical representation of it. The obvious choice is;

$$\eta_{ab} \to g_{ab}(x), \quad \partial_a \to \nabla_a = \partial_a + \Gamma_a, \tag{2.1}$$

where Γ_a is a frame-dependent object.

This way the so called *metricity* requirements, $\nabla_a g_{bc} = 0$, follow from the trivial flat space relations, $\partial_a \eta_{bc} = 0$. If one defines the action of covariant derivative on the metric tensor by,

$$\nabla_a g_{bc} = \partial_a g_{bc} - \Gamma^e_{ab}\, g_{ec} - \Gamma^e_{ac}\, g_{eb},$$

then the *torsionless* condition, $\Gamma_{ab} = \Gamma_{ba}$, can be used to solve the metricity conditions. So, the principle of general covariance represents a technical way to express the transition from flat space Lorentzian physics to the curved space of GR. Notice that we have already used the principle of general covariance when we assumed that $\mathcal{L}(x)$ is some unspecified scalar.

Ok, we are ready to do the job. First notice that $g_{ab}(x)$ would obey a second order differential equation if $\mathcal{L}(x)$ were a function of g_{ab} and $\partial_c g_{ab}$ only. But, the principle of equivalence makes it impossible to have a non-trivial scalar function $\mathcal{L}(g_{ab}, \partial_c g_{ab})$; any such function can be made equal to the constant $\mathcal{L}(\eta_{ab}, 0)$, because it is always possible to set $g_{ab} = \eta_{ab}$ and $g_{ab,c} = 0$ at any point by coordinate transformation. The only way out of this is to let $\mathcal{L}$ to depend on g_{ab} and its frist ($g_{ab,c}$) and second ($g_{ab,cd}$) derivatives but demand that $\frac{\partial \mathcal{L}}{\partial g_{ab,cd}}$ be a function of g_{ab} only. According to Eq (1.9), $g_{ab}(x)$ will then satisfy a second order differential equation. So, we may write $\mathcal{L}$ in the form;

$$\mathcal{L}(g_{ab}, g_{ab,c}, g_{ab,cd}) = g_{ab,cd}(x)\, A^{abcd}(g_{ab}) + B(g_{ab}, g_{ab,c}). \tag{2.2}$$

Let us evaluate this Lagrangian in a locally inertial system; At the point $x^a = 0$, we choose coordinates such that $g_{ab}(0) = \eta_{ab}$ and $g_{ab,c}(0) = 0$. Hence,

$$\mathcal{L}\left(\eta_{ab}, 0, g_{ab,cd}(x)\right) = g_{ab,cd}(x)\, A^{abcd}(\eta_{ab}) + b, \tag{2.3}$$

where $b = B(\eta_{ab}, 0)$. In a new coordinate system related to the x-system by the Lorentz transformation $x^a = \Lambda^a{}_b \bar{x}^b$, we still have $\bar{g}_{ab} = \eta_{ab}$, and $\bar{g}_{ab,c} = 0$, but

$$\bar{g}_{ab,cd} = \Lambda^m{}_a\, \Lambda^n{}_b\, \Lambda^p{}_c\, \Lambda^q{}_d\, g_{mn,pq}. \tag{2.4}$$

Since $\mathcal{L}$ is a Lorentz scalar, we must have

$$\bar{g}_{ab,cd}\, A^{abcd} = g_{ab,cd}\, A^{abcd}. \tag{2.5}$$

Eq (2.4) and Eq (2.5) imply that $A^{abcd}(\eta_{nm})$ is an invariant Lorentz tensor. In flat space-time, the most general rank-4 invariant tensor is

$$A^{abcd} = a\eta^{ab}\, \eta^{cd} + a_1\, \eta^{ac}\, \eta^{bd} + a_2\, \eta^{ad}\, \eta^{bc} + a_3\, \epsilon^{abcd}. \tag{2.6}$$

Using the symmetry of $g_{ab,cd}$ in a and b and in c and d, we can write Eq(2.3) in the form

$$\mathcal{L} = g_{ab,cd}\, \left(a\eta^{ab}\eta^{cd} + c\eta^{ac}\eta^{bd}\right) + b, \tag{2.7}$$

where a and c are constants. Next, we go to yet another (locally inertial) coordinate system related to the x-system by

$$x^a = \bar{x}^a + \frac{1}{6}\,\eta^{ae}\, C_{ebcd}\bar{x}^b\,\bar{x}^c\,\bar{x}^d, \tag{2.8}$$

where the constant C_{ebcd} is symmetric in b, c and d. With some boring calculation (see proof below), at $\bar{x}^a = 0$, one finds that,

$$\bar{g}_{ab,cd} = g_{ab,cd} + C_{abcd} + C_{bacd}. \tag{2.9}$$

Now, it is easy to see that the invariance of $\mathcal{L}(x)$ under this transformation;

$$\mathcal{L}(\eta_{ab}, 0, \bar{g}_{ab,cd}) = \mathcal{L}(\eta_{ab}, 0, g_{ab,cd}),$$

implies $a = -c$. Thus, our Lagrangian becomes

$$\mathcal{L}(x) = c g_{ab,cd}(x)\,\left(\eta^{ac}\,\eta^{bd} - \eta^{ab}\,\eta^{cd}\right) + b. \tag{2.10}$$

Now comes the most difficult part, we play with the indices and rewrite the first term in (2.10) as

$$\frac{c}{2}\,\eta^{bc}\,\eta^{ae}\,\partial_a\left(g_{be,c} + g_{ce,b} - g_{bc,e}\right) - \frac{c}{2}\,\eta^{ae}\,\eta^{bc}\,\partial_c\left(g_{be,a} + g_{ae,b} - g_{ba,e}\right), \tag{2.11}$$

and, after introducing the “connection coefficients”,

$$\hat{\Gamma}^a_{bc} = \frac{1}{2}\,\eta^{ae}\,(g_{ce,b} + g_{be,c} - g_{bc,e}),$$

we conclude that in a locally inertial frame our Lagrangian has the form

$$\mathcal{L}(x) = c\,\eta^{bc}\left(\partial_a\hat{\Gamma}^a_{bc} - \partial_c\hat{\Gamma}^a_{ba}\right) + b. \tag{2.12}$$

Finally, we are in good position to use the principle of general covariance;(2.1), and write an expression for $\mathcal{L}(x)$ which holds true in any, *completely arbitrary*, reference frame,

$$\mathcal{L}(x) = c\,R + b, \tag{2.13}$$

where,

$$R = g^{ab}\,\left(\nabla_c\Gamma^c_{ab} - \nabla_b\Gamma^c_{ac}\right), \tag{2.14}$$

is the scalar curvature and[1],

$$\Gamma^a_{bc} = \frac{1}{2}g^{ae}(g_{ce,b} + g_{be,c} - g_{bc,e}).$$

The values of c and b must be determined by experiment. For $b = 0$, the value of c can be determined by comparing the Newtonian limit of the theory with Newtons gravity, this gives $c = \frac{1}{16\pi G}$. Thus, for $b = \frac{\Lambda}{16} \neq 0$, we arrive at the H-E action[2]:

$$S[g_{ab}] = \frac{1}{16\pi G} \int d^4x \sqrt{-g}\, (R + \Lambda)\,. \tag{2.15}$$

3 Conclusion

We have shown that the H-E action follows from the two principles of general relativity and the assumption that the field satisfies a second order diferential equation. Thus, there must be a room for this assumption in the set of axioms of general relativity. We have also seen that the principle of general covariance played a crucial part in the derivation. So, in spite of all the fuss about its "physical content", general covariance seems to be a powerful and important principle which determines acceptable forms of physical laws.

.1 Proof of equation (2.9)

By differentiating the coordinate transformation, (2.8), and substituing the result in the right hand side of (1.5) we find

$$\bar{g}_{ab}(\bar{x}) = g_{ab}(x) + \frac{1}{2}\, g_{ae}\, \eta^{en}\, C_{nbpq} \bar{x}^q\, \bar{x}^p + \frac{1}{2}\, g_{eb}\, \eta^{en}\, C_{napq}\, \bar{x}^p\, \bar{x}^q + O(\bar{x}^4),$$

Now we differentiate with respect to $\bar{x}^c$, we find,

$$\bar{g}_{ab,c} = g_{ab,c} + g_{ae}\, \eta^{en}\, C_{nbpc}\, \bar{x}^p + g_{eb}\, \eta^{en}\, C_{napc}\, \bar{x}^p + F,$$

where F represents a collection of some junk that will vanish when we evaluate the final expression at $\bar{x}^a = 0$;

$$F = O(\bar{x}^2)\, \bar{\partial} g + \cdots$$

Now we take the second derivative and collect more junk;

$$\bar{g}_{ab,cd} = g_{ab,cd} + g_{ae}\, \eta^{en}\, C_{nbcd} + g_{eb}\, \eta^{en}\, C_{nacd} + G(\bar{x}),$$

where G represents some thing of order;

$$G = \bar{\partial} F + O(\bar{x})\, \bar{\partial} g$$

At $\bar{x} = 0$ we can put $g_{ab} = \eta_{ab}$, and since the junk G(0) vanish, we find our result

$$\bar{g}_{ab,cd} = g_{ab,cd} + \delta^n_a\, C_{nbcd} + \delta^n_b\, C_{nacd}.$$

Exercise. consider the linear transformation,

$$x^a = \bar{x}^a + \eta^{an}\, C_{nm}\, \bar{x}^m.$$

Show that,

$$\bar{g}_{ab} = g_{ab} + C_{ab} + C_{ba}.$$

Now, consider

$$x^a = \bar{x}^a + \frac{1}{2}\,\eta^{an}\,C_{npq}\,\bar{x}^p\,\bar{x}^q,$$

then show, at $\bar{x} = 0$, that

$$\bar{g}_{ab,c} = g_{ab,c} + C_{abc} + C_{bac}.$$

References

[1] S. Weinberg, *Gravitation and Cosmology,* Wiley, New York, 1972.

[2] D. Hilbert, Math. Ann. 92, 1 (1924).