The least action principle for the Einstein field equation

Foster Teng
APM426H1 General Relativity
Instructor: R. McCann

March 27, 2025

Abstract

This paper presents a detailed and rigorous derivation of the Einstein field equation using variational calculus. By defining the Einstein-Hilbert action. we start from varying the gravitational action by applying some essential identities including the Palatini identity, Jacobi's formula, and so on, which eventually gives us the Einstein equation in vacuum. Next, we define the matter action, whose variation gives the definition of the stress–energy tensor. In the end, we add variation of these two actions together to derive the full Einstein field equation.

Contents

1 Introduction

In classical mechanics, we define the Lagrangian as the kinetic energy minus the potential energy and define the action by taking the integral of the Lagrangian, then we apply the variation to derive the equation of motion by assuming the extremum of the action. In general relativity, we apply a similar approach by defining the Einstein-Hilbert action and derive the Einstein field equation. This method provides another approach and insight to the derivation of the Einstein field equation instead of starting from the postulates. Since the Einstein-Hilbert action consists of the Ricci scalar and the metric determinant, we apply the variation separately to each one of them. The most part of this paper is to rigorously prove some useful identities and apply the divergence theorem, properties of determinant functions, and careful treatment of variation of the metric on the boundary to do the variation on the Ricci scalar. Next, the variation of the metric determinant is easily derived using the previously proved identity. By combining the action of matter as well, we finally derive the full Einstein field equation.

2 The Einstein–Hilbert Action and the Variational Principle

We begin by introducing the Einstein–Hilbert action, which is the action functional for the gravitational field in general relativity. It is defined as:

$$
\begin{aligned}
S &= S_{\text{gravitational}} + S_{\text{matter}} \qquad (1)\\
&= \frac{1}{2\kappa}\int R\sqrt{-g}\,d^4x + S_{\text{matter}}. \qquad (2)
\end{aligned}
$$

where $\kappa = 8\pi G$, g is the determinant of the metric tensor $g_{\mu\nu}$, R is the Ricci scalar, $S_{\text{gravitational}} = \frac{1}{2\kappa}\int R\sqrt{-g}\,d^4x$, and S_{matter} is the action for any matter fields present.

We aim to derive the Einstein's field equations by applying the principle of stationary action:

$$\delta S = 0. \tag{3}$$

First we need to apply the variation on $S_{\text{gravitational}}$ with respect to the metric tensor $g^{\mu\nu}$ by expanding $\delta(R\sqrt{-g})$ as

$$\delta(R\sqrt{-g}) = (\delta R)\sqrt{-g} + R(\delta\sqrt{-g}) \tag{4}$$

Secondly, to also vary R, we write $R = R_{\mu\nu}g^{\mu\nu}$ to further vary $R_{\mu\nu}$ and $g^{\mu\nu}$ as well.

Finally, we will include the variation on S_{matter} together to derive the full Einstein field equation.

3 Variation of the Ricci Scalar

In this section, we apply the variation on the Ricci Scalar in the first term of Equation (4). Before deriving it, we consider some identities, and the first is the Palatini identity.

3.1 Proof of the Palatini Identity[2]: $\delta R_{\mu\nu} = \nabla_\lambda \delta\Gamma^\lambda_{\mu\nu} - \nabla_\nu \delta\Gamma^\lambda_{\mu\lambda}$

The Palatini identity expresses the variation of the Ricci tensor $R_{\mu\nu}$ in terms of the variation of the Christoffel symbols $\delta\Gamma^\lambda_{\mu\nu}$:

$$\delta R_{\mu\nu} = \nabla_\lambda \delta\Gamma^\lambda_{\mu\nu} - \nabla_\nu \delta\Gamma^\lambda_{\mu\lambda}. \tag{5}$$

To derive this identity, we begin with the definition of the Riemann curvature tensor in terms of the Levi-Civita connection:

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}. \tag{6}$$

Taking the variation of both sides, we obtain:

$$\delta R^\rho{}_{\sigma\mu\nu} = \partial_\mu \delta\Gamma^\rho_{\nu\sigma} - \partial_\nu \delta\Gamma^\rho_{\mu\sigma} + \delta\Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} + \Gamma^\rho_{\mu\lambda}\delta\Gamma^\lambda_{\nu\sigma} - \delta\Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} - \Gamma^\rho_{\nu\lambda}\delta\Gamma^\lambda_{\mu\sigma}.$$

Although Christoffel symbols $\Gamma^\rho_{\nu\sigma}$ themselves are not tensors, the variation $\delta\Gamma^\rho_{\nu\sigma}$ is a tensor. Therefore, we are allowed to take its covariant derivative:

$$\nabla_\mu \delta\Gamma^\rho_{\nu\sigma} = \partial_\mu \delta\Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\mu\lambda}\delta\Gamma^\lambda_{\nu\sigma} - \Gamma^\lambda_{\mu\nu}\delta\Gamma^\rho_{\lambda\sigma} - \Gamma^\lambda_{\mu\sigma}\delta\Gamma^\rho_{\nu\lambda}. \tag{7}$$

Solving this equation for $\partial_\mu \delta\Gamma^\rho_{\nu\sigma}$ and substituting it into the variation of the Riemann tensor, we find that all terms involving $\Gamma\delta\Gamma$ cancel out. Therefore, we are left with the following.

$$\delta R^\rho{}_{\sigma\mu\nu} = \nabla_\mu \delta\Gamma^\rho_{\nu\sigma} - \nabla_\nu \delta\Gamma^\rho_{\mu\sigma}, \tag{8}$$

We now contract the first and third indices to obtain the variation of the Ricci tensor:

$$\begin{aligned} \delta R_{\nu\sigma} &= \delta R^\mu{}_{\nu\mu\sigma} && (9)\\ &= \nabla_\mu \delta\Gamma^\mu_{\sigma\nu} - \nabla_\sigma \delta\Gamma^\mu_{\mu\nu}, && (10) \end{aligned}$$

which is the desired Palatini identity, and it will be used in the next subsection.

3.2 A little rearrangement by applying Palatini Identity

We then consider the variation of the Ricci scalar R with respect to the metric. The Ricci scalar is given by $R = g^{\mu\nu}R_{\mu\nu}$, where $R_{\mu\nu}$ is the Ricci tensor, and its variation is:

$$\begin{aligned} \delta R &= \delta(g^{\mu\nu}R_{\mu\nu}) && (11)\\ &= R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu}. && (12) \end{aligned}$$

Then by the Palatini identity.

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\left(\nabla_\lambda \delta\Gamma^\lambda_{\mu\nu} - \nabla_\nu \delta\Gamma^\lambda_{\mu\lambda}\right). \tag{13}$$

Using the linearity and product rule of the covariant derivative, we can write:

$$\begin{aligned} g^{\mu\nu}\nabla_\lambda \delta\Gamma^\lambda_{\mu\nu} &= \nabla_\lambda(g^{\mu\nu}\delta\Gamma^\lambda_{\mu\nu}) - (\nabla_\lambda g^{\mu\nu})\delta\Gamma^\lambda_{\mu\nu}, && (14)\\ g^{\mu\nu}\nabla_\nu \delta\Gamma^\lambda_{\mu\lambda} &= \nabla_\nu(g^{\mu\nu}\delta\Gamma^\lambda_{\mu\lambda}) - (\nabla_\nu g^{\mu\nu})\delta\Gamma^\lambda_{\mu\lambda}. && (15) \end{aligned}$$

Because the Levi-Civita connection is metric compatible ($\nabla_\alpha g^{\mu\nu} = 0$), the second terms in both expressions vanish, so Equation (13) becomes:

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\lambda\left(g^{\mu\nu}\delta\Gamma^\lambda_{\mu\nu} - g^{\mu\lambda}\delta\Gamma^\nu_{\mu\nu}\right). \tag{16}$$

define

$$V^\lambda = g^{\mu\nu}\delta\Gamma^\lambda_{\mu\nu} - g^{\mu\lambda}\delta\Gamma^\nu_{\mu\nu},$$

So Equation (16) becomes:

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\lambda V^\lambda \tag{17}$$

3.3 Proof of the Jacobi's Formula[2]: $\partial_\mu g = g\, g^{\alpha\beta}\, \partial_\mu g_{\beta\alpha}$

We are going to argue that the second term $\nabla_\lambda V^\lambda$ of Equation (17) vanishes when we apply the variation of it to the integral. But before doing so, we first need to prove the Jacobi formula.

3.3.1 Identification of Determinant Function

Let us identify the determinant as a real-valued function:

$$\det : \mathbb{R}^{n^2} \to \mathbb{R}$$

where $\mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$, and identify the directional derivative of the determinant evaluated at

$$A \in \mathbb{R}^{n^2}$$

along the direction

$$T \in \mathbb{R}^{n^2}$$

as:

$$\det{}'(A)(T) = \frac{d}{d\varepsilon} \det(A + \varepsilon T)\bigg|_{\varepsilon=0}.$$

for a general point

$$A \in \mathbb{R}^{n^2}$$

, but here we require A to be invertible!

3.3.2 Lemma 1

$$\det{}'(I)(T) = \mathrm{tr}(T).$$

Proof. Using the definition of the directional derivative, we compute:

$$\det{}'(I)(T) = \frac{d}{d\varepsilon} \det(I + \varepsilon T)\bigg|_{\varepsilon=0}.$$

Since $\det(I + \varepsilon T)$ is a polynomial of degree n in ε, its linear term is $\mathrm{tr}(T)$, so:

$$\det{}'(I)(T) = \mathrm{tr}(T).$$

Therefore, we have:

$$\begin{aligned}
\det{}'(I)(T) &= \frac{d}{d\varepsilon} \det(I + \varepsilon T)\bigg|_{\varepsilon=0} \\
&= \lim_{\varepsilon\to 0} \frac{\det(I + \varepsilon T) - \det I}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{1 + \mathrm{tr}(T)\varepsilon + \mathcal{O}(\varepsilon^2) - \det I}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{\mathrm{tr}(T)\varepsilon + \mathcal{O}(\varepsilon^2)}{\varepsilon} \\
&= \mathrm{tr}(T)
\end{aligned}$$

□

3.3.3 Lemma 2

$$\det'(A)(T) = \det A \cdot \mathrm{tr}(A^{-1}T).$$

Proof. By property of the determinant of matrices, we have:

$$\det(X) = \det(A) \cdot \det\big(A^{-1}X\big).$$

Differentiating the determinant function at $X = A$ along the direction T, we get:

$$\begin{aligned}
\det'(A)(T) &= \frac{d}{d\varepsilon}\det(A+\varepsilon T)\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\big(\det(A)\det\big(I+\varepsilon A^{-1}T\big)\big)\Big|_{\varepsilon=0} \\
&= \det(A)\frac{d}{d\varepsilon}\det\big(I+\varepsilon A^{-1}T\big)\Big|_{\varepsilon=0} \\
&= \det A \cdot \det'(I)(A^{-1}T) \quad \text{(by definition)} \\
&= \det A \cdot \mathrm{tr}(A^{-1}T) \quad \text{(by lemma 1)}
\end{aligned}$$

□

3.3.4 Proof of the Jacobi's Formula[3]

Let $A(t)$ depending on t be a differentiable family of invertible $n \times n$ matrices. Then:

$$\frac{d}{dt}\det A(t) = \mathrm{tr}\left(A^{-1} \cdot \frac{dA}{dt}\right).$$

Proof. setting $T = \frac{dA}{dt}$, we have:

$$\det'(A)(T) = \det'(A) \cdot \frac{dA}{dt} = \frac{d}{dt}\det A$$

where the middle term is a row vector multiplied by a column vector, and both are with n*n tuples

$$\det A \cdot \mathrm{tr}(A^{-1}T) = \det A \cdot \mathrm{tr}\left(A^{-1}\frac{dA}{dt}\right).$$

Then both equations are equal by Lemma 2

□

Therefore, with the identification $g = det(A)$, $\partial_\mu g = \frac{d}{dt}\det A(t)$, $g^{\alpha\beta} = A^{-1}$, $g_{\beta\alpha} = A$, and $\partial_\mu g_{\beta\alpha} = \frac{dA}{dt}$, we get $\partial_\mu g = g\, g^{\alpha\beta}\, \partial_\mu g_{\beta\alpha}$.

3.4 Proof of the Identity[1]: $\sqrt{-g}\nabla_\lambda V^\lambda = \partial_\lambda(\sqrt{-g}V^\lambda)$

The covariant derivative of V^α is given by:

$$\nabla_\alpha V^\alpha = \partial_\alpha V^\alpha + \Gamma^\alpha_{\mu\alpha}V^\mu$$

where

$$\Gamma^{\alpha}_{\mu\alpha} = \frac{1}{2} g^{\alpha\beta} \left(\partial_\mu g_{\beta\alpha} + \partial_\alpha g_{\beta\mu} - \partial_\beta g_{\mu\alpha} \right)$$

Because the second and the third terms are antisymmetric in the indices α and β, only the first term survives:

$$\Gamma^{\alpha}_{\mu\alpha} = \frac{1}{2} g^{\alpha\beta} \partial_\mu g_{\beta\alpha}$$

Therefore,

$$\begin{aligned}
\sqrt{-g}\nabla_\alpha V^\alpha &= \sqrt{-g}\partial_\alpha V^\alpha + \sqrt{-g}\Gamma^{\alpha}_{\mu\alpha} V^\mu \\
&= \sqrt{-g}\partial_\mu V^\mu + \sqrt{-g}\frac{1}{2} g^{\alpha\beta}\partial_\mu g_{\beta\alpha} V^\mu \quad \text{(by the equality we just proved above)} \\
&= \sqrt{-g}\partial_\mu V^\mu + \sqrt{-g}\frac{1}{2}\frac{\partial_\mu g}{g} V^\mu \quad \text{(by the Jacobi's Formula)} \\
&= \sqrt{-g}\partial_\mu V^\mu + \sqrt{-g}\frac{1}{2}\frac{\partial_\mu(-g)}{(-g)} V^\mu \\
&= \sqrt{-g}\partial_\mu V^\mu + \frac{1}{2}\frac{\partial_\mu(-g)}{\sqrt{-g}} V^\mu \\
&= \sqrt{-g}\partial_\mu V^\mu + (\partial_\mu\sqrt{-g}) V^\mu \\
&= \partial_\alpha(\sqrt{-g} V^\alpha)
\end{aligned}$$

□

3.5 Divergence as a Total Derivative and Boundary Term

Now, after a long journey, if we come back to the variation, by the identity we proved in the subsection 3.4. In the integral, the second term of Equation (17) becomes:

$$\int_{\mathcal{M}} \sqrt{-g}\, \nabla_\lambda V^\lambda \, d^4x = \int_{\mathcal{M}} \partial_\lambda(\sqrt{-g} V^\lambda)\, d^4x. \tag{18}$$

By the divergence theorem (Stokes' theorem in differential form),

$$\int_{\mathcal{M}} \partial_\lambda(\sqrt{-g} V^\lambda)\, d^4x = \int_{\partial\mathcal{M}} \sqrt{h}\, V^\lambda n_\lambda \, d^3x, \tag{19}$$

where n_λ is the outward-pointing unit normal to the boundary $\partial\mathcal{M}$ and $\sqrt{h}$ is the induced volume element on the boundary, i.e. it is the determinant of the restricted metric g on the boundary of M.

In general, the integrand of this surface integral is not guaranteed to vanish but recall that,

$$V^\lambda = g^{\mu\nu}\delta\Gamma^{\lambda}_{\mu\nu} - g^{\mu\lambda}\delta\Gamma^{\nu}_{\mu\nu}$$

The variation on the christoffel symbols involves the term

$$\delta g^{\mu\nu}$$

and its derivative. If the spacetime manifold M has the boundary, then we have to assume that they are zero to ensure that this variational problem is well-posed, otherwise we have to add an extra term, which is called Gibbons–Hawking–York boundary term , in our action in order to cancel them out. So this is in fact an important assumption in this paper. Physically, it means that we only consider the events in the interior of M and ignore the variation of geometry on the boundary. Finally, if the spacetime manifold has no boundary, then the surface integral vanishes since the trivial boundary has measure zero. Therefore, we have:

$$\int_{\mathcal{M}} \sqrt{-g}\, \nabla_\lambda V^\lambda \, d^4x = 0. \tag{20}$$

Hence, the term

$$\nabla_\lambda V^\lambda$$

eventually does not contribute to our variational problem. We are done with the variation of the first term in Equation (4), and the conclusion is the following:

$$\int_{\mathcal{M}} (\delta R)\sqrt{-g} d^4x = \int_{\mathcal{M}} (R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\lambda V^\lambda)\sqrt{-g} d^4x \quad \text{(by Equation (17))} \tag{21}$$

$$= \int_{\mathcal{M}} R_{\mu\nu}\delta g^{\mu\nu}\sqrt{-g} d^4x \quad \text{(by Equation (20))} \tag{22}$$

4 Variation of the Metric Determinant

In this section, we apply the variation on the metric determinant in the second term of Equation (4). This part is easy. We just compute:

$$\begin{aligned}
\delta\sqrt{-g} &= \frac{1}{2}\frac{\delta(-g)}{\sqrt{-g}} \\
&= \frac{1}{2}\frac{\delta(-g)}{-g}\sqrt{-g} \\
&= \frac{1}{2} g^{\mu\nu}\delta g_{\mu\nu}\sqrt{-g} \quad \text{(by Jacobi's formula)} \\
&= -\frac{1}{2} g_{\mu\nu}\delta g^{\mu\nu}\sqrt{-g} \quad (\text{since } g_{\mu\nu}\delta g^{\mu\nu} = -g^{\mu\nu}\delta g_{\mu\nu})
\end{aligned}$$

Therefore,

$$\int_{\mathcal{M}} R(\delta\sqrt{-g}) d^4x = \int_{\mathcal{M}} -\frac{1}{2} R g_{\mu\nu}\delta g^{\mu\nu}\sqrt{-g} d^4x \tag{23}$$

5 Combining Variations to Derive the Field Equations

Compute the variation of the gravitational part of the Einstein–Hilbert action:

$$\delta S_{\text{gravitational}} = \frac{1}{2\kappa}\int_{\mathcal{M}} \delta(R\sqrt{-g}) d^4x \tag{24}$$

$$= \frac{1}{2\kappa}\int_{\mathcal{M}} \left[(\delta R)\sqrt{-g} + R(\delta\sqrt{-g})\right] d^4x \quad \text{(by Equation (4))} \tag{25}$$

$$= \frac{1}{2\kappa}\int_{\mathcal{M}} (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu})\delta g^{\mu\nu}\sqrt{-g} d^4x \quad \text{(by Equation (22) and (23))} \tag{26}$$

$\delta S = 0$ for arbitrary $\delta g^{\mu\nu}$ yields the vacuum Einstein field equations:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0. \tag{27}$$

6 Inclusion of Matter Fields

We now consider the presence of matter by including a matter action S_{matter} defined by:

$$S_{\text{matter}} = \int_{\mathcal{M}} \mathcal{L}_{\text{M}} \sqrt{-g} d^4x \tag{28}$$

The variation of this action with respect to the metric is as follows:

$$\delta S_{\text{matter}} = \int_{\mathcal{M}} \delta(\mathcal{L}_{\text{M}} \sqrt{-g}) d^4x \tag{29}$$

$$= \frac{-1}{2} \int_{\mathcal{M}} (-2) \frac{\delta(\mathcal{L}_{\text{M}} \sqrt{-g})}{\sqrt{-g}\delta g^{\mu\nu}} \delta g^{\mu\nu} \sqrt{-g} d^4x \tag{30}$$

define the stress-energy tensor:

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta(\mathcal{L}_{\text{M}} \sqrt{-g})}{\delta g^{\mu\nu}} \tag{31}$$

So Equation (30) becomes:

$$\delta S_{\text{matter}} = \frac{-1}{2} \int_{\mathcal{M}} T_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x \tag{32}$$

Recall that the total action is $S = S_{\text{gravitational}} + S_{\text{matter}}$, so $\delta S = 0$ implies that:

$$0 = \frac{1}{2\kappa} \int_{\mathcal{M}} (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}) \delta g^{\mu\nu} \sqrt{-g} d^4x + \frac{-1}{2} \int_{\mathcal{M}} T_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x \quad \text{(by Equation (26) and (32))} \tag{33}$$

$$= \int_{\mathcal{M}} \left[\frac{1}{2\kappa} (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}) - \frac{1}{2} T_{\mu\nu} \right] \delta g^{\mu\nu} \sqrt{-g} d^4x \tag{34}$$

which leads to the full Einstein field equations:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \kappa T_{\mu\nu}. \tag{35}$$

for κ being a very small constant

7 Conclusion

Through a careful and rigorous application of the variational principle to the Einstein–Hilbert action, we have derived Einstein field equations. The process involved multiple non-trivial identities, including the Palatini identity and Jacobi formula, as well as handling boundary terms via the divergence theorem. The variation of both the Ricci scalar and the metric determinant was computed explicitly. The inclusion of matter fields through the matter Lagrangian completed the derivation of the full Einstein equations. This work demonstrates not only the mathematical consistency of general relativity, but also the elegance of the least action principle as a foundational tool in gravitational theory.

References

[1] B. F. Schutz. *A First Course in General Relativity*, 2nd ed., Cambridge University Press, 2009. See §6.3, pp. 152–153.

[2] Wikipedia contributors. *Einstein–Hilbert action*. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Einstein%E2%80%93Hilbert_action. Accessed March 26, 2025.

[3] MIT OpenCourseWare. *8.962 General Relativity, Spring 2020*. YouTube playlist. https://www.youtube.com/playlist?list=PLUl4u3cNGP629n_3fX7HmKKgin_rqGzbx. Accessed March 26, 2025.