4 General Theory of Relativity

Physical theories in mathematical form are complete, if mathematical structure is identified with physical observables. In general relativity, the spacetime is a four-dimensional manifold with a metric, which can be measured by local experiments, and the metric can be locally made Minkowski by a choice of coordinate. Furthermore, freely falling particles move on a time-like geodesic (massless particles on a null geodesic). In addition to this identification, we need to derive the Einstein equation, describing how physical objects and a curved spacetime are related.

4.1 Einstein Equations

4.1.1 Principle of General Covariance

The principle of general covariance states that the physical laws should take the same (tensorial) form in all coordinate systems:

- The equations in general should reduce to those in special relativity (in the absence of gravity), when $g_{\mu\nu} \rightarrow \eta_{ab}$ and $\Gamma^{\rho}_{\mu\nu} \rightarrow 0$,
- the equations should be in a covariant form.

This implies that we can derive the general tensorial equations for most physical laws that are valid in the presence of gravity, simply by taking those in special relativity (free-falling system) without gravity and making them covariant:

$$\eta_{ab} \quad \rightarrow \quad g_{\mu\nu} \,, \qquad\qquad m\frac{du^a}{d\tau} = f^a \quad \rightarrow \quad m\frac{Du^\mu}{d\tau} = f^\mu \,, \tag{4.1}$$

where the covariant derivative of the four velocity is

$$\frac{Du^\mu}{d\tau} := \frac{du^\mu}{d\tau} + \Gamma^\mu_{\rho\sigma} u^\rho u^\sigma = u^\nu u^\mu{}_{;\nu} \,. \tag{4.2}$$

In the same way, the Maxwell's equations (1.84) and (1.85) in special relativity becomes

$$\partial_\nu F^{\mu\nu} = 4\pi J^\mu \quad \rightarrow \quad \nabla_\nu F^{\mu\nu} = 4\pi J^\mu \,, \qquad\qquad 0 = \varepsilon^{\rho\sigma\mu\nu}\partial_\sigma F_{\mu\nu} \quad \rightarrow \quad 0 = \varepsilon^{\rho\sigma\mu\nu}\nabla_\sigma F_{\mu\nu} \,. \tag{4.3}$$

4.1.2 Energy-Momentum Tensor for Macroscopic Fluids

For our purposes, we are not interested in the microscopic states of the systems, but interested in their macroscopic states, often described by the density, the pressure, the temperature, and so on. Since the energy-momentum tensor was already derived in special relativity, we can use the principle of general covariance to generalize the energy-momentum tensor in a curved spacetime.

The energy-momentum tensor for a fluid can be expressed in terms of the fluid quantities measured by an observer with four velocity u^μ as (*the most general decomposition*)

$$T_{\mu\nu} := \rho u_\mu u_\nu + p\mathcal{H}_{\mu\nu} + q_\mu u_\nu + q_\nu u_\mu + \pi_{\mu\nu} \,, \qquad\qquad 0 = \mathcal{H}_{\mu\nu}u^\nu \,, \tag{4.4}$$

where $\mathcal{H}_{\mu\nu}$ is the projection tensor and

$$\mathcal{H}_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu \,, \qquad \mathcal{H}^\mu_\mu = 3 \,, \qquad u^\mu q_\mu = 0 = u^\mu \pi_{\mu\nu} \,, \qquad \pi_{\mu\nu} = \pi_{\nu\mu} \,, \qquad \pi^\mu_\mu = 0 \,. \tag{4.5}$$

The variables ρ, p, q_μ and $\pi_{\mu\nu}$ are the energy density, the isotropic pressure (including the entropic one), the (spatial) energy flux and the anisotropic pressure measured by the observer with u_μ, respectively, i.e.,

$$\rho = T_{\mu\nu}u^\mu u^\nu \,, \qquad p = \frac{1}{3}T_{\mu\nu}\mathcal{H}^{\mu\nu} \,, \qquad q_\mu = -T_{\rho\sigma}u^\rho \mathcal{H}^\sigma_\mu \,, \qquad \pi_{\mu\nu} = T_{\rho\sigma}\mathcal{H}^\rho_\mu \mathcal{H}^\sigma_\nu - p\mathcal{H}_{\mu\nu} \,. \tag{4.6}$$

Remember that these fluid quantities are observer-dependent.

4.1.3 Einstein's Equation

While most of the physical laws can be deduced by using the principle of general covariance, the physical laws for gravity cannot be deduced, as there is no gravity in a free-falling system, where special relativity is valid. In the non-relativistic limit, the Newtonian mechanics works, and the gravitational field equation should reduce to the Poisson equation

$$\nabla^2\phi = 4\pi G\rho \,. \tag{4.7}$$

Given that the metric and the energy momentum tensor in the Newtonian limit are

$$g_{00} \simeq -\left(1+\frac{2\phi}{c^2}\right) , \qquad T_{00} \simeq \rho c^2 \,, \tag{4.8}$$

the gravitational field equation in the Newtonian limit is expected to be

$$\nabla^2 g_{00} = -\frac{8\pi G}{c^4} T_{00} \,, \tag{4.9}$$

where the proportionality constant was set by the Poisson equation $\Delta\phi = 4\pi G\rho$. Note that the left-hand side is a function of metric, and the right hand-side is a matter component.

Using the principle of general covariance, we can deduce the gravitational field equation in any coordinate to take a tensorial form:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \,, \qquad \lim_{\rm nr} G_{\mu\nu} = -\delta^\mu_0\delta^\nu_0\nabla^2 g_{\mu\nu} \,, \tag{4.10}$$

where the limit represents the non-relativistic case and $G_{\mu\nu}$ is a function of metric. The requirement for the LHS is as follows: $G_{\mu\nu}$ should be tensorial, and it should be symmetric and conserved:

$$G_{\mu\nu} = G_{(\mu\nu)} \,, \qquad G_{\mu\nu;\mu} = 0 \,, \tag{4.11}$$

as $T_{\mu\nu}$ is symmetric and conserved. Furthermore, the dimension of $G_{\mu\nu}$ should be two in mass dimension:

$$[T_{\mu\nu}] = L^{-4} \,, \qquad 8\pi G = \frac{1}{M^2_{\rm pl}} = L^2 \,, \qquad \therefore \;\; [G_{\mu\nu}] = L^{-2} \,. \tag{4.12}$$

A tensor $G_{\mu\nu}$ that is a function of metric tensor $g_{\mu\nu}$ and of mass dimension two is uniquely determined up to two arbitrary dimensionless constants as

$$G_{\mu\nu} = c_1 R_{\mu\nu} + c_2 R\, g_{\mu\nu} \,, \tag{4.13}$$

which is already symmetric. The conservation condition $G_{\mu\nu;\mu} = 0$ imposes

$$c_2 = -\frac{1}{2}c_1 \,, \qquad G_{\mu\nu} \equiv c_1\left(R_{\mu\nu} - \frac{1}{2}R\, g_{\mu\nu}\right) = \frac{8\pi G}{c^4}\, T_{\mu\nu} \,. \tag{4.14}$$

The remaining constant c_1 can be determined by taking the non-relativistic limit:

$$G_{00} = c_1\left(R_{00} - \frac{1}{2}R\, \eta_{00}\right) \gg G_{\alpha\beta} \,, \tag{4.15}$$

where the inequality arises from the non-relativistic conditions in $T_{\mu\nu}$. Noting that in this limit

$$g^{\mu\nu}G_{\mu\nu} = -c_1 R \simeq \eta^{00}G_{00} = -G_{00} \,, \tag{4.16}$$

we just need to compute only one component R_{00} of the Riemann tensor or just the Ricci scalar R:

$$G_{00} \simeq 2c_1 R_{00} \simeq c_1 R \,, \qquad R \simeq 2R_{00} \,. \tag{4.17}$$

The Ricci tensor in the non-relativistic limit is

$$R_{\mu\nu} = R^\rho{}_{\mu\rho\nu} \simeq \Gamma^\rho_{\mu\nu,\rho} - \Gamma^\rho_{\mu\rho,\nu} \,, \qquad R_{00} \simeq \Gamma^i_{00,i} \simeq -\frac{1}{2}\Delta h_{00} \,, \tag{4.18}$$

where we kept only the spatial derivative terms. Therefore, we obtain the Einstein tensor G_{00} and fix the coefficient c_1:

$$G_{00} \simeq -c_1\nabla^2 h_{00} \,, \qquad c_1 \equiv 1 \,. \tag{4.19}$$

The gravitational field equation (or Einstein's equations) is

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}R\, g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \,, \qquad R = -\frac{8\pi G}{c^4}\, T \,. \tag{4.20}$$

Given the energy-momentum tensor, the Ricci scalar is algebraically determined.

4.1.4 Structure of the Einstein Field Equation

The Einstein equations $G_{\mu\nu} = 8\pi G\, T_{\mu\nu}$ constrain the relation between the metric and the energy momentum tensor. In a Machian view, the existence of matter determines the spacetime geometry, and the matter moves along the straightest path in a given spacetime geometry. This is summarized in a well-known statement by John Wheeler: "matter tells how to curve, and spacetime tells matter how to move."

• ***Degrees of freedom***.— It is clear that any symmetric tensor $g_{\mu\nu}$ plugged in the Einstein equations will result in unphysical energy-momentum tensor. In general, the field equation is considered as "constraints" for twenty independent variables in $g_{\mu\nu}$ and $T_{\mu\nu}$. Given ten dofs in $T_{\mu\nu}$, ten Einstein equations provide ten constraints for $g_{\mu\nu}$. Due to the Bianchi identity, however, four of the Einstein equations are redundant. This situation of under-determination is saved by general covariance or gauge transformation, which takes away four dofs in $g_{\mu\nu}$ or $T_{\mu\nu}$.

• ***Nonlinearity***.— The Einstein equations are non-linear, such that one cannot add two independent solutions to derive another solution. In other words, one cannot analyze a complex system by breaking it into several simpler systems and adding the solutions of the simpler systems to obtain the solution of the complex system. Consider a matter source, and it produces gravitational fields, which contain energy. This energy of the gravitational fields is then equivalent to mass, which in turn produces gravitational fields. Due to this complexity in the field equation, Einstein anticipated that no one would find an exact solution to the nonlinear equations, and he was greatly surprised to see Schwarzschild solution, less than a year after general relativity was published.

• ***Hidden boundary conditions?***.— Einstein was puzzled by the existence of a Minkowski solution for an empty universe,

$$g_{\mu\nu} \equiv \eta_{\mu\nu} \, , \qquad\qquad R_{\mu\nu} = T_{\mu\nu} = 0 \, , \tag{4.21}$$

which would be in conflict with Mach's principle. Some unknown boundary conditions might be needed to exclude such solutions. Einstein in the end added a cosmological constant Λ to the field equation:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda\, g_{\mu\nu} = 8\pi G\, T_{\mu\nu} \, , \tag{4.22}$$

to remove the Minkowski solution and to make the Universe static. After the introduction of his cosmological constant in 1917, Edwin Hubble found in 1929 that the Universe was expanding and Einstein allegedly said "it was my biggest blunder!" A cosmological constant makes the Universe unstable and indeed expand in an accelerating manner. Furthermore, de Sitter also found a solution in an empty universe with non-zero cosmological constant.

• ***Geodesic equation without field equation***.— The particle motion or the geodesic equation is expected to follow from the Einstein field equation, as the matter determines the spacetime and the spacetime affects the trajectory. It turns out, however, that the geodesic equation is rather generic and follows only from the conservation equation, *not* from the field equation. Consider dust particles ρ moving with a time-like vector u^μ, and the conservation equation yields

$$T_{\mu\nu} = \rho u_\mu u_\nu \, , \qquad\qquad 0 = \nabla_\nu T^{\mu\nu} = \rho u^\nu \nabla_\nu u^\mu + u^\mu \nabla_\nu (\rho u^\nu) \, . \tag{4.23}$$

Given the normalization $-1 = u_\mu u^\mu$, and $0 = u_\mu \nabla_\nu u^\mu$, the conservation equation yields

$$0 = \nabla_\nu (\rho u^\nu) \, , \qquad\qquad 0 = u^\nu \nabla_\nu u^\mu \, . \tag{4.24}$$

The latter is the geodesic equation.

4.1.5 Cauchy Problem: Initial Value Problem

The Cauchy problem refers to the case of whether the physical system given the governing equation can be solved from the initial conditions. In general relativity, the initial conditions are the metric tensor $g_{\mu\nu}$ at some initial time t (or hypersurface) and its time derivative $g_{\mu\nu,0}$. Mind that given the information on the initial hypersurface, the spatial derivative of the metric tensor is also available, i.e., all first derivative $g_{\mu\nu,\rho}$ information at the initial hypersurface.

For simplicity, we consider a vacuum solution to the Einstein equation $R_{\mu\nu} = 0$. If we obtain the condition for $g_{\mu\nu,00}$ from the Einstein equation and the initial conditions, we obtain all the derivatives of the metric tensor by subsequently taking derivatives of the Einstein equation and we derive the solution for $g_{\mu\nu}$ everywhere all the time, provided that the metric tensor is differentiable and analytic.

In fact, computation of the Einstein equation shows

$$R_{00} = -\frac{1}{2}g^{\alpha\beta}g_{\alpha\beta,00} + I_{00} = 0 \,, \qquad R_{0\alpha} = \frac{1}{2}g^{0\beta}g_{\alpha\beta,00} + I_{0\alpha} = 0 \,, \qquad R_{\alpha\beta} = -\frac{1}{2}g^{00}g_{\alpha\beta,00} + I_{\alpha\beta} = 0 \,, \tag{4.25}$$

where $I_{\mu\nu}$ are the tensors written only in terms of the initial conditions. The dynamical equations contain no information about $g_{00,00}$ or $g_{0\alpha,00}$ (*under-determination*), or ten dynamical equations *over-determine* six $g_{\alpha\beta,00}$. The first issue can be removed by a specific coordinate transformation

$$\tilde{x}^\mu = x^\mu + \frac{1}{6}t^3F^\mu(x) \,, \tag{4.26}$$

to set $\tilde{g}_{00,00} = \tilde{g}_{0\alpha,00} = 0$, while keeping the initial conditions unchanged:

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} \,, \qquad \tilde{g}_{\mu\nu,\rho} = g_{\mu\nu,\rho} \,, \qquad \tilde{g}_{\mu\nu,\alpha\rho} = g_{\mu\nu,\alpha\rho} \,, \tag{4.27}$$

but

$$\tilde{g}_{00,00} = g_{00,00} - 2g_{0\mu}F^\mu \,, \qquad \tilde{g}_{0\alpha,00} = g_{0\alpha,00} - g_{\alpha\mu}F^\mu \,, \qquad \tilde{g}_{\alpha\beta,00} = g_{\alpha\beta,00} \,. \tag{4.28}$$

The dynamical equations are indeed *not* over-determined. Provided $g^{00} \neq 0$, the last dynamical equation $R_{\alpha\beta} = 0$ can be used to solve for $g_{\alpha\beta,00}$, then two remaining dynamical equations are then the constraint equations on the initial conditions. Re-arranging the dynamical equations, we obtain

$$g^{00}R_{00} - g^{\alpha\beta}R_{\alpha\beta} = g^{00}I_{00} - g^{\alpha\beta}I_{\alpha\beta} = 0 \,, \qquad g^{00}R_{0\alpha} + g^{0\beta}R_{\alpha\beta} = g^{00}I_{0\alpha} + g^{0\beta}I_{\alpha\beta} = 0 \,, \tag{4.29}$$

which are also equivalent to

$$G_\alpha{}^0 = 0 \,. \tag{4.30}$$

As long as the main dynamical equation $R_{\alpha\beta} = 0$ is satisfied, $R_{00} = 0$ and $R_{0\alpha} = 0$ are satisfied only in terms of initial conditions. Hence, the Einstein equations are now composed of six dynamical equations $R_{\alpha\beta} = 0$ and four constraint equations $G_\alpha{}^0 = 0$. Once the constraint equations are satisfied on the initial hypersurface, they are satisfied all the time due to the Bianchi identity:

$$0 = \nabla_\nu G_\mu{}^\nu = \nabla_0 G_\mu{}^0 + \nabla_\alpha G_\mu{}^\alpha \,, \qquad \nabla_0 G_\mu{}^0 = -\nabla_\alpha G_\mu{}^\alpha = 0 \,, \tag{4.31}$$

where the second equality can be proved by showing that $\nabla_\alpha G_\mu{}^\alpha$ is a linear combination of $G_\mu{}^0$ and $G_\mu{}^0{}_{,\alpha}$, both of which vanish on the initial hypersurface.

• also add § 8 in Wald (1984).

4.2 Einstein-Hilbert Action

4.2.1 Gauge Transformation

The general covariance of general relativity guarantees that any coordinate system can be used to describe the physics and it has to be independent of coordinate systems. This is known as the diffeomorphism symmetry in general relativity. However, when we split the metric into the background and the perturbations around it by choosing a coordinate system, we explicitly change the correspondence of the physical Universe to the background homogeneous and isotropic Universe. Hence, the metric perturbations transform non-trivially (or gauge transform), and the diffeomorphism invariance implies that the physics should be gauge-invariant.

The gauge group of general relativity is the group of diffeomorphisms. A diffeomorphism corresponds to a differentiable coordinate transformation. The coordinate transformation on the manifold $\mathcal{M}$ can be considered as one generated by a smooth vector field ζ^μ. Given the vector field ζ^μ, consider the solution of the differential equation

$$\left.\frac{d\chi^\mu(\lambda)}{d\lambda}\right|_P = \zeta^\mu\left[\chi^\nu_P(\lambda)\right] \,, \qquad \chi^\mu_P(\lambda = 0) = x^\mu_P \,, \qquad \frac{d}{d\lambda} = \zeta^\mu\partial_\mu \,, \tag{4.32}$$

defines the parametrized integral curve $x^\mu(\lambda) = \chi^\mu_P(\lambda)$ with the tangent vector $\zeta^\mu(x_P)$ at P. Therefore, given the vector field ζ^μ on $\mathcal{M}$ we can define an associated coordinate transformation on $\mathcal{M}$ as $x^\mu_P \to \tilde{x}^\mu_P = \chi^\mu_P(\lambda = 1)$ for any given P. Assuming that ζ^μ is small one can use the perturbative expansion for the solution of equation to obtain

$$\begin{aligned}\tilde{x}^\mu_P &= \chi^\mu_P(\lambda=1) = \chi^\mu_P(\lambda=0) + \frac{d}{d\lambda}\chi^\mu_P\bigg|_{\lambda=0} + \frac{1}{2}\frac{d^2}{d\lambda^2}\chi^\mu_P\bigg|_{\lambda=0} + \cdots \\ &= x^\mu_P + \zeta^\mu(x_P) + \frac{1}{2}\zeta^\mu{}_{,\nu}\zeta^\nu + \mathcal{O}(\zeta^3) = e^{\zeta^\nu\partial_\nu}x^\mu \,. \end{aligned} \tag{4.33}$$

This parametrization corresponds to the gauge-transformation with ζ^μ.

In general, any gauge-transformation of tensor $\mathbf{T}$ for an infinitesimal change ζ can be expressed in terms of the Lie derivative (valid to all orders of $\mathbf{T}$)

$$\delta_\zeta \mathbf{T} := \tilde{\mathbf{T}} - \mathbf{T} = -\pounds_\zeta \mathbf{T} + \mathcal{O}(\zeta^2) \,, \tag{4.34}$$

To all orders in ζ, we have

$$\tilde{\mathbf{T}}(x) = \mathbf{T}(x) - \pounds_\zeta\mathbf{T} + \frac{1}{2}\pounds^2_\zeta\mathbf{T} + \cdots = \exp\left[-\pounds_\zeta\right]\mathbf{T} \,. \tag{4.35}$$

Therefore, the gauge-transformation in perturbation theory is simply

$$\delta_\zeta\bar{\mathbf{T}} = 0 \,, \qquad \delta_\zeta\mathbf{T}^{(1)} = -\pounds_\zeta\bar{\mathbf{T}} \,, \qquad \delta_\zeta\mathbf{T}^{(n)} = -\pounds_\zeta\mathbf{T}^{(n-1)} \,, \tag{4.36}$$

where we used that ζ is also a perturbation. At the linear order, the Lie derivative is trivial, and the the most general coordinate transformation in Eq. (4.34) becomes

$$\tilde{x}^\mu = x^\mu + \xi^\mu \,, \tag{4.37}$$

where we now use $\xi^\mu = \zeta^\mu$. The transformation of the metric tensor at the leading order in ξ is then

$$\delta_\xi g_{\mu\nu}(x) := \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = -\pounds_\xi g_{\mu\nu} = -\left(\xi_{\mu;\nu} + \xi_{\nu;\mu}\right) \,, \tag{4.38}$$

where the semi-colon represents the covariant derivative with respect to the full metric $g_{\mu\nu}$.

4.2.2 Variational Method and Equation of Motion

Physical theories can be concisely expressed in terms of its action S, and the principle of least action states that the equation of motion in a given theory follows the path that minimizes the change in S. The action S is a general and non-local function of its dynamic variables such as ϕ, ψ, A_μ, $g_{\mu\nu}$, but in most cases the actions can be expressed in terms of *local functions*, called Lagrangian:

$$S =: \int dt\, L =: \int d^4x\, \mathcal{L} \,, \tag{4.39}$$

where L is the Lagrangian and $\mathcal{L}$ is the Lagrangian (density). The dimension

$$[S] = \frac{M\cdot L^2}{T} \sim h \,, \qquad [L] = \frac{M\cdot L^2}{T^2} \sim E \,, \qquad [\mathcal{L}] = \frac{M}{L\cdot T^2} \sim \rho_E \,, \tag{4.40}$$

where h is the Planck constant whose dimension is the same as the action as the angular momentum (times the angle) or the energy times the time.

Since our interest is a theory of gravity, the Lagrangian should be a function of metric $g_{\mu\nu}$ and its spacetime derivatives as

$$\mathcal{L}_g = \mathcal{L}_g(g_{\mu\nu},\, g_{\mu\nu,\rho},\, g_{\mu\nu,\rho\sigma},\, \cdots) \,. \tag{4.41}$$

When we make a variation to the dynamical variable $\tilde{g}_{\mu\nu} = g_{\mu\nu} + \delta g_{\mu\nu}$, the action must remain stationary:

$$0 = \tilde{S} - S = \int d^4x\, \frac{\delta\mathcal{L}_g}{\delta g_{\mu\nu}}\,\delta g_{\mu\nu} \,, \qquad \mathcal{L}^{\mu\nu}_g := \frac{\delta\mathcal{L}_g}{\delta g_{\mu\nu}} = 0 \,, \tag{4.42}$$

and we have derived the equation of motion for $\mathcal{L}_g$. Note that the tensor $\mathcal{L}^{\mu\nu}$ is symmetric and of density -1, as $g_{\mu\nu}$ is symmetric and the volume factor is of weight $+1$. Furthermore, the general covariance guarantees that we can use any coordinates, and under a general coordinate transformation the metric tensor transforms as in Eq. (4.38). In this case, the least action condition yields

$$\delta g_{\mu\nu} = -2\,\xi_{(\mu;\nu)}\ , \qquad 0 = \delta S = -2\int d^4x\ \mathcal{L}_g^{\mu\nu}\nabla_\nu\xi_\mu = -2\int d^4x\Big[\nabla_\nu(\mathcal{L}^{\mu\nu}\xi_\mu) - \xi_\mu\nabla_\nu\mathcal{L}^{\mu\nu}\Big]\ . \tag{4.43}$$

The first term in the square bracket vanishes, because a divergence of a tensor of weight -1 can be replaced with a coordinate derivative and integration by part yields a surface term, where assume that the metric variation $\delta g_{\mu\nu}$ vanishes on the boundary of the manifold. Hence, we obtain another condition for our gravity theory $\mathcal{L}_g$

$$\nabla_\nu\mathcal{L}^{\mu\nu} = 0\ . \tag{4.44}$$

In fact, these conditions are valid for any gravity theories, as long as it is a local function of metric tensor and it contains the general covariance.

4.2.3 Einstein-Hilbert Action

Having discussed the general properties of $\mathcal{L}_g$, we look for a specific Lagrangian for general relativity. First, note that $\mathcal{L}_g$ should be also a scalar of weight -1. The simplest possibility we can construct out of $g_{\mu\nu}$ is

$$\mathcal{L}_g \equiv \sqrt{-g}\ , \tag{4.45}$$

and we can readily derive

$$\delta(\sqrt{-g}) = \frac{1}{2}\sqrt{-g}\ g^{\mu\nu}\delta g_{\mu\nu}\ , \qquad \mathcal{L}^{\mu\nu} = \frac{1}{2}\sqrt{-g}\ g^{\mu\nu}\ , \qquad \nabla_\nu\mathcal{L}^{\mu\nu} = 0\ . \tag{4.46}$$

This theory of gravity is not dynamic. The next simplest possibility as a function of $g_{\mu\nu}$ is called the *Einstein-Hilbert* action:

$$\mathcal{L}_{\rm EH} := \sqrt{-g}\ R\ , \qquad S_{\rm EH} := \int d^4x\ \sqrt{-g}\ \frac{c^4}{16\pi G}\ R\ , \tag{4.47}$$

where we restored the dimensionful constants in the action. The variation of the Einstein-Hilbert action gives two terms

$$\delta(\sqrt{-g}\ R) = \frac{1}{2}\sqrt{-g}\ g^{\mu\nu}\delta g_{\mu\nu}R + \sqrt{-g}\ \delta R\ , \tag{4.48}$$

and we need to compute the variation of the Ricci tensor

$$R = g^{\mu\nu}R_{\mu\nu}\ , \qquad \delta R = \delta g^{\mu\nu}R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu}\ , \tag{4.49}$$

where the first term can be manipulated by using $\delta^\mu_\nu = g^{\mu\rho}g_{\rho\nu}$ as

$$g^{\mu\rho}\delta g_{\rho\nu} = -g_{\rho\nu}\delta g^{\mu\rho}\ , \qquad \delta g^{\mu\nu} = -g^{\mu\rho}g^{\nu\sigma}\delta g_{\rho\sigma}\ , \qquad \delta g_{\mu\nu} = -g_{\mu\rho}g_{\nu\sigma}\delta g^{\rho\sigma}\ , \tag{4.50}$$

and the second term can be from the Palatini identity in Eq. (3.122):

$$\delta R_{\mu\nu} = \left(\delta\Gamma^\rho_{\mu\nu}\right)_{;\rho} - \left(\delta\Gamma^\rho_{\mu\rho}\right)_{;\nu}\ . \tag{4.51}$$

Foremost, we deal with the second term with $\delta R_{\mu\nu}$ and show that it is just a surface term in the action. With $\nabla_\rho g_{\mu\nu} = 0$, we first compute

$$g^{\mu\nu}\delta R_{\mu\nu} = \left(g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu}\right)_{;\rho} - \left(g^{\mu\nu}\delta\Gamma^\rho_{\mu\rho}\right)_{;\nu} = \nabla_\mu\left(g^{\rho\sigma}\delta\Gamma^\mu_{\rho\sigma} - g^{\rho\mu}\delta\Gamma^\sigma_{\rho\sigma}\right) =: \nabla_\mu V^\mu\ , \tag{4.52}$$

where we re-arranged the dummy indicies and defined a vector V^μ. The contribution of this term in the action is then

$$\sqrt{-g}\ \nabla_\mu V^\mu = \frac{\partial}{\partial x^\mu}\left(\sqrt{-g}\ V^\mu\right)\ , \tag{4.53}$$

integrated by part and removed. The remaining terms in the action is now

$$\delta S_{\rm EH} = \frac{c^4}{16\pi G}\int d^4x\,\sqrt{-g}\,\delta g_{\mu\nu}\left[\frac{1}{2}g^{\mu\nu}R - g^{\rho\mu}g^{\sigma\nu}R_{\rho\sigma}\right] =: \int d^4x\,\delta g_{\mu\nu}\,\mathcal{L}^{\mu\nu}_{\rm EH}\,, \tag{4.54}$$

and the Einstein (vacuum) field equation is

$$0 = -\mathcal{L}^{\mu\nu}_{\rm EH} = R^{\mu\nu} - \frac{1}{2}\,R\,g^{\mu\nu} =: G^{\mu\nu}\,, \tag{4.55}$$

where we defined the Einstein tensor $G_{\mu\nu}$. Given the general covariance, we derived the differential identity:

$$0 = \nabla_\nu\mathcal{L}^{\mu\nu}_{\rm EH} = \nabla_\nu G^{\mu\nu}\,, \tag{4.56}$$

which is known as the *Bianchi identity*.

In the Einstein-Hilbert action, we can add a cosmological constant term in terms of the simplest Lagrangian $\mathcal{L}_g$ we considered first in Eq. (4.45)

$$\mathcal{L}_\Lambda := -\frac{\Lambda}{8\pi G}\sqrt{-g}\,, \qquad S_g = \int d^4x\,\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{\Lambda}{8\pi G}\right)\,, \tag{4.57}$$

and this additional term yields

$$0 = -\mathcal{L}^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}\,R\,g^{\mu\nu} + \Lambda\,g^{\mu\nu}\,. \tag{4.58}$$

4.2.4 Matter Action and Energy-Momentum Tensor

Now we consider the matter Lagrangian $\mathcal{L}_m$ in particle physics, which includes scalar fields, vector fields (E&M), and other fluids. With the matter action, the full action including gravity becomes

$$S = S_g + S_m = \int d^4x\,\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{\Lambda}{8\pi G} + \mathcal{L}_m\right)\,, \tag{4.59}$$

where we kept the notation for the matter Lagrangian $\mathcal{L}_m$.[1] The variation of S_m with respect to the metric yields

$$\delta S_m = \int d^4x\sqrt{-g}\,\delta g_{\mu\nu}\left(\frac{1}{2}\,g^{\mu\nu}\mathcal{L}_m + \frac{\delta\mathcal{L}_m}{\delta g_{\mu\nu}}\right)\,, \tag{4.60}$$

and the full Einstein field equations becomes

$$G^{\mu\nu} = \frac{16\pi G}{c^4}\left(\frac{1}{2}\,g^{\mu\nu}\mathcal{L}_m + \frac{\delta\mathcal{L}_m}{\delta g_{\mu\nu}}\right) =: \frac{8\pi G}{c^4}T^{\mu\nu}\,, \tag{4.61}$$

where we defined the energy-momentum tensor in a formal way as

$$T^{\mu\nu} := \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g_{\mu\nu}} = g^{\mu\nu}\mathcal{L}_m + 2\,\frac{\delta}{\delta g_{\mu\nu}}\int d^4x\,\mathcal{L}_m\,. \tag{4.62}$$

Given that we can define

$$\delta S_m = \int d^4x\,\delta g_{\mu\nu}\mathcal{L}^{\mu\nu}_m\,, \qquad \mathcal{L}^{\mu\nu}_m = \frac{1}{2}\,\sqrt{-g}\,T^{\mu\nu}\,, \tag{4.63}$$

the differential identity condition gives the energy-momentum conservation

$$0 = \nabla_\nu\mathcal{L}^{\mu\nu}_m = \nabla_\nu T^{\mu\nu}\,. \tag{4.64}$$

For the same computation, we obtain

$$T_{\mu\nu} := -\frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}} = g_{\mu\nu}\mathcal{L}_m - 2\,\frac{\delta}{\delta g^{\mu\nu}}\int d^4x\,\mathcal{L}_m\,. \tag{4.65}$$

[1] Mind that the Lagrangian density we defined is a scalar of weight -1, while in particle physics $\mathcal{L}_m$ is derived in the Minkowski spacetime $\sqrt{-\eta}\equiv 1$. So we kept $\mathcal{L}_m$ but add $\sqrt{-g}$ to the Lagrangian in general relativity.

Mind the subtlety, for example, for the scalar field action,

$$\mathcal{L}_\phi = -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \,, \qquad \delta\mathcal{L}_\phi = -\frac{1}{2}\delta g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi = \frac{1}{2}\delta g_{\mu\nu}\partial^\mu\phi\,\partial^\nu\phi \,. \tag{4.66}$$

• ***E&M Action***.— The action for the E&M is

$$S_{\rm E\&M} = \int d^4x\,\sqrt{-g}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + 4\pi\,J^\mu A_\mu\right) \,. \tag{4.67}$$

Using

$$F_{\mu\nu}F^{\mu\nu} = g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma} \,, \qquad (\partial_\rho A_\sigma)\,F_{\mu\nu} = \partial_\rho\left(A_\sigma F_{\mu\nu}\right) - A_\sigma\partial_\rho F_{\mu\nu} \,, \tag{4.68}$$

we vary the action with respect to A_λ and perform the integration by part to derive the equation of motion

$$\nabla_\nu F^{\mu\nu} = 4\pi J^\mu \,. \tag{4.69}$$

Mind that the integration by part is performed by a coordinate derivative but the metric tensor commutes only with a covariant derivative, which is possible due to the density weight of the Lagrangian. By varying the action with respect to the metric tensor, we then obtain the energy-momentum tensor for E&M:

$$T_{\mu\nu} = F_{\mu\rho}F_\nu{}^\rho + g_{\mu\nu}\left(-\frac{1}{4}F_{\rho\sigma}F^{\rho\sigma} + 4\pi J^\rho A_\rho\right) \,. \tag{4.70}$$

4.2.5 Structure of the Einstein-Hilbert Action

The Lagrangian for the Einstein-Hilbert action is a local function of the metric tensor at a given spacetime, and it contains the second derivatives of $g_{\mu\nu}$ in the Ricci scalar (first derivative in the Christoffel symbol). While the full action is a functional of the metric tensor field $g_{\mu\nu}$, the (local) Lagrangian is just a function of $g_{\mu\nu}$, first derivative $g_{\mu\nu,\rho}$, and second derivative $g_{\mu\nu,\rho\sigma}$. Hence, the variation with respect to $g_{\mu\nu}$ yields

$$\delta\mathcal{L}_{\rm EH} = \frac{\partial\mathcal{L}}{\partial g_{\mu\nu}}\delta g_{\mu\nu} + \frac{\partial\mathcal{L}}{\partial g_{\mu\nu,\rho}}\delta g_{\mu\nu,\rho} + \frac{\partial\mathcal{L}}{\partial g_{\mu\nu,\rho\sigma}}\delta g_{\mu\nu,\rho\sigma} + \mathcal{O}(\delta g^2) \,, \tag{4.71}$$

and a few integrations by part then give the equation of motion

$$\mathcal{L}^{\mu\nu}_{\rm EH} = \frac{\partial\mathcal{L}}{\partial g_{\mu\nu}} - \frac{\partial}{\partial x^\rho}\frac{\partial\mathcal{L}}{\partial g_{\mu\nu,\rho}} + \frac{\partial^2}{\partial x^\rho\partial x^\sigma}\frac{\partial\mathcal{L}}{\partial g_{\mu\nu,\rho\sigma}} \,. \tag{4.72}$$

Given the last term with four derivatives, it is surprising that the Einstein equation in the end contains only the second derivatives of $g_{\mu\nu}$. Indeed, an explicit computation yields

$$\frac{\partial\mathcal{L}}{\partial g_{\mu\nu,\rho\sigma}} = \sqrt{-g}\left[\frac{1}{2}\left(g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}\right) - g^{\mu\nu}g^{\rho\sigma}\right] \,. \tag{4.73}$$

This feature is unique to general relativity.

Ostrogradski (1850) showed that a *non-degenerate* Lagrangian with finite higher-order derivative variables (than ordinary one derivative) leads to a unbounded Hamiltonian, i.e., the theory includes a dynamical variable with negative kinetic energy (or Ostrogradski *ghost*). The Lagrangian $\mathcal{L}_{\rm EH}$ is indeed degenerate:

$$\mathcal{H}_{ij} := \frac{\partial^2\mathcal{L}_{\rm EH}}{\partial\dot{q}^i\,\partial\dot{q}^j} \,, \qquad \det\,\mathcal{H}_{ij} = 0 \,, \tag{4.74}$$

or the lapse N and the shift N^i are Lagrange multipliers, where i, j here indicate the dynamical variables.

4.2.6 Palatini Method

In a curved spacetime, there are two essential ingredients that determine the inner product and the parallel transport, i.e., the metric tensor $g_{\mu\nu}$ and the affine connection $\Gamma^\rho_{\mu\nu}$. One can think of these two objects as two independent spacetime fields, while they are related to each other in general relativity. In fact, the Einstein-Hilbert action can be exclusively written in terms of these two elements:

$$\mathcal{L}_{\rm EH} = \sqrt{-g}\, g^{\mu\nu} R_{\mu\nu} = \sqrt{-g}\, g^{\mu\nu} \left(\Gamma^\epsilon_{\mu\nu,\epsilon} - \Gamma^\epsilon_{\mu\epsilon,\nu} + \Gamma^\epsilon_{\mu\nu}\Gamma^\rho_{\rho\epsilon} - \Gamma^\epsilon_{\mu\rho}\Gamma^\rho_{\nu\epsilon}\right) = \mathcal{L}_{\rm EH}(g_{\mu\nu}, \Gamma^\rho_{\mu\nu}, \Gamma^\rho_{\mu\nu,\sigma}) \,, \tag{4.75}$$

arising from the fact that the Ricci tensor is just a function of $\Gamma^\rho_{\mu\nu}$ and its first derivatives.

Treating the affine connection as an independent dynamic variable, we apply the variational method with respect to $\Gamma^\rho_{\mu\nu}$ and derive

$$\delta\mathcal{L}_{\rm EH} = \sqrt{-g}\, g^{\mu\nu}\delta R_{\mu\nu} = \sqrt{-g}\, g^{\mu\nu}\left[\left(\delta\Gamma^\rho_{\mu\nu}\right)_{;\rho} - \left(\delta\Gamma^\rho_{\mu\rho}\right)_{;\nu}\right] , \tag{4.76}$$

where we again used the Palatini identity. Mind that $g_{\mu\nu}$ is *not* varied, as it is independent from $\Gamma^\rho_{\mu\nu}$ in this approach. An integration by part yields

$$\delta\mathcal{L}_{\rm EH} = -\nabla_\rho\left(\sqrt{-g}\, g^{\mu\nu}\right)\delta\Gamma^\rho_{\mu\nu} + \nabla_\nu\left(\sqrt{-g}\, g^{\mu\nu}\right)\delta\Gamma^\rho_{\mu\rho} = \left[\delta^\nu_\rho\nabla_\sigma\left(\sqrt{-g}\, g^{\mu\sigma}\right) - \nabla_\rho\left(\sqrt{-g}\, g^{\mu\nu}\right)\right]\delta\Gamma^\rho_{\mu\nu} \,, \tag{4.77}$$

where in the second equality we re-arranged the dummy indicies. The least action principle states that the square bracket vanishes against the variation $\delta\Gamma^\rho_{\mu\nu}$. Note, however, that since $\delta\Gamma^\rho_{\mu\nu}$ is symmetric over μ,ν, only the symmetric part of the square bracket should vanish:

$$0 = \delta^\nu_\rho\nabla_\sigma\left(\sqrt{-g}\, g^{\mu\sigma}\right) + \delta^\mu_\rho\nabla_\sigma\left(\sqrt{-g}\, g^{\nu\sigma}\right) - 2\nabla_\rho\left(\sqrt{-g}\, g^{\mu\nu}\right) \,, \tag{4.78}$$

where we made the first term in $\delta\mathcal{L}_{\rm EH}$ symmetric over μ,ν.

The variation with respect to the affine connection gives a differential equation for the metric tensor, and we can show that this equation is satisfied, only if the metricity condition is satisfied

$$0 = \nabla_\rho\, g_{\mu\nu} = \partial_\rho\, g_{\mu\nu} - \Gamma^\epsilon_{\rho\mu}\, g_{\epsilon\nu} - \Gamma^\epsilon_{\rho\nu}\, g_{\mu\epsilon} \,. \tag{4.79}$$

In other words, the Palatini approach treats the affine connection as an independent variable, and the equation of motion for the connection gives the metricity condition, which then leads to the relation to the metric tensor, i.e., Christoffel symbol in general relativity. Note that the metricity condition in the Palatini approach is a consequence, not a choice.

It is also clear from the above derivation that in any other theories of gravity the relation between the affine connection and the metric tensor will not be the same, when the Palatini approach is applied.