ENS M1 General Relativity - 8 - The Gravitational Action Principle

Nick Kaiser

November 19, 2020

Contents

1 Introduction

The route we have taken in developing GR followed that charted by Einstein. We assert that gravitational phenomena are the influence of curvature of the space-time manifold. Requiring agreement with Newtonian theory we are led – aside from possible ambiguities associate with the cosmological constant – to a unique rank-two contraction $\mathbf{G}$ of the curvature tensor that is 'sourced' by the matter stress-tensor $\mathbf{T}$.

A radically different way of thinking about the field equations is that they are the 'Euler-Lagrange' equations that emerge from the requirement that the *Einstein-Hilbert action*

$$\boxed{S = \int d^4x \sqrt{g} \left(\frac{R}{16\pi\kappa} + \mathcal{L}_{\rm m} \right)} \tag{1}$$

where R is the Ricci scalar – a function of the metric $\mathbf{g}$ and its derivatives – and where $\mathcal{L}_{\rm m}$ is the Lorentz scalar Lagrangian density for the matter fields, denoted loosely by $\phi(\vec{x})$, be extremised with respect to variation of the metric.

So just as the usual action principle states that, given some space-time $\mathbf{g}$, the matter field configurations $\phi(\vec{x})$ that actually occur in nature are those that extremise the action for the matter, the *gravitational action principle* states that the space-times that actually occur in nature are those that extremise the complete action given above.

As well as being profound, this is practically useful for two reasons. One is that it is a good starting point for thinking about possible modifications to Einstein's gravity (beyond just adding a cosmological term $\Lambda\mathbf{g}$ to $\mathbf{G}$). The other is that this leads to a stress-energy tensor $T^{\mu\nu}$ for the matter that is guaranteed to be symmetric, whereas, as we saw earlier, in electromagnetism for instance, this was not the case.

We will now confirm that this does indeed lead to Einstein's equations. To do this we will first show in §3 that the variation of the part of the action involving the gravitational field is

$$\delta S_{\rm g} \equiv \delta \int d^4x \sqrt{g} \frac{R}{16\pi\kappa} = -\frac{1}{16\pi\kappa} \int d^4x \sqrt{g} G^{\mu\nu} \delta g_{\mu\nu} \tag{2}$$

where $G^{\mu\nu}$ is the Einstein tensor.

Then, in §4, we will show that

$$\delta(\sqrt{g}\mathcal{L}_{\rm m}) = \tfrac{1}{2}\sqrt{g} T^{\mu\nu} \delta g_{\mu\nu} \tag{3}$$

so the stress energy tensor $\mathbf{T}$ can be considered to be the functional derivative of the matter Lagrangian density with respect to the metric. Together, these establish that $\mathbf{G} = 8\pi\kappa\mathbf{T}$ are indeed the Euler-Lagrange equations that emerge from $\delta S = 0$.

First, however, we will remind ourselves how the energy and momentum conservation laws $\vec{\nabla}\cdot\mathbf{T} = 0$ follow via Noether's theorem from invariance of the matter Lagrangian density $\mathcal{L}_{\rm m}$ with respect to translations in (flat) space-time.

2 The stress-energy tensor for the matter in flat space-time

2.1 The Noether currents for matter fields in flat space-time

Let us assume that there is some Lagrangian density for the 'matter fields' – matter being defined loosely here as it may contain massless radiation – that is the function of the fields and their derivatives and, perhaps, position:

$$\mathcal{L}(\phi, \phi_{,\alpha}, \vec{x}) \tag{4}$$

where we have written the field as a scalar quantity, but it is really shorthand here for a collection of fields that will also in general contain things like the electromagnetic field A^μ.

If we assume that this is Lorentz invariant – so we allow products of different fields to give interactions[1] provided they are also Lorentz scalars – we get a Lorentz invariant action – space-time volume element d^4x being Lorentz invariant –

$$S = \int d^4x \mathcal{L}(\phi, \phi_{,\alpha}, \vec{x}). \tag{5}$$

The equations of motion are obtained by requiring that the action be stationary with respect to variations of the fields, so $\delta S = \int d^4x \delta\mathcal{L} = 0$. The variation of the Lagrangian density is

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_{,\nu}}\delta\phi_{,\nu} + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi \\ &= \left(\frac{\partial\mathcal{L}}{\partial\phi_{,\nu}}\delta\phi\right)_{,\nu} + \delta\phi\left[-\left(\frac{\partial\mathcal{L}}{\partial\phi_{,\nu}}\right)_{,\nu} + \frac{\partial\mathcal{L}}{\partial\phi}\right] \end{aligned} \tag{6}$$

where we have used $\partial_\nu\delta\phi = \delta\phi_{,\nu}$ and have 'differentiated by parts' to eliminate $\delta\phi_{,\nu}$. It is still there, of course, hidden in the first term. But this is a total derivative, so it does not contribute to δS if, as usual, we demand that the field be fixed on the boundary (or if we assume that the fields tend to zero at infinite, or if we assume there is no boundary – as in a closed universe).

For δS to vanish for an arbitrary $\delta\phi$ requires that the quantity $[\ldots]$ above vanish, which gives the Euler-Lagrange equations – there being one for each field, or field component –

$$\boxed{\frac{\partial}{\partial x^\nu}\frac{\partial\mathcal{L}}{\partial\phi_{,\nu}} = \frac{\partial\mathcal{L}}{\partial\phi}.} \tag{7}$$

If we write $\mathcal{L}(\vec{x}) = \mathcal{L}(\phi(\vec{x}), \phi_{,\alpha}(\vec{x}), \vec{x})$ for some actual solution of the field equations, and take its partial derivative with respect the $\mu^{\rm th}$ components of $\vec{x}$, keeping the others fixed, we have

$$\partial_\mu\mathcal{L}(\vec{x}) = \underbrace{\frac{\partial\mathcal{L}}{\partial\phi}\phi_{,\mu} + \frac{\partial\mathcal{L}}{\partial\phi_{,\nu}}\phi_{,\nu\mu}}_{\partial_\nu(\phi_{,\mu}\partial\mathcal{L}/\partial\phi_{,\nu})} + \partial_\mu\mathcal{L} \tag{8}$$

where we have invoked the equations of motion to eliminate $\partial\mathcal{L}/\partial\phi$ and where $\mathcal{L}$ on the right hand side is being considered to be $\mathcal{L}(\phi, \phi_{,\alpha}, \vec{x})$.

Writing the left hand side as $\partial_\mu\mathcal{L} = \delta^\nu_\mu\partial_\nu\mathcal{L}$, and raising the index μ, gives

$$T^{\nu\mu}{}_{,\nu} = \partial^\mu\mathcal{L}(\phi, \phi_{,\alpha}, \vec{x}) \tag{9}$$

where the stress-energy tensor is

$$\boxed{T^{\nu\mu} \equiv -\phi^{,\mu}\frac{\partial\mathcal{L}}{\partial\phi_{,\nu}} + \eta^{\nu\mu}\mathcal{L}.} \tag{10}$$

[1] In electromagnetism this is done by replacing the ordinary operator ∂_μ by the gauge-covariant derivative $D_\mu \equiv \partial_\mu + (q/\hbar)A_\mu$.

If the Lagrangian density has no explicit dependence on the space-time coordinates – i.e. $\mathcal{L}(\phi, \phi_{,\alpha}, \vec{x}) = \mathcal{L}(\phi, \phi_{,\alpha})$ – then the right hand side of (9) vanishes and we have four continuity equations (one for each of $\mu = 0, 1, 2, 3$ and expressing continuity of energy and the three components of momentum)

$$\boxed{T^{\nu\mu}{}_{,\nu} = 0} \tag{11}$$

or, if you prefer,

$$\boxed{\vec{\nabla} \cdot \mathbf{T} = 0}. \tag{12}$$

We say that the symmetry of the Lagrangian density $\mathcal{L}$ with respect to translations in each if the 4-dimensions of space-time has given rise to 4 conserved 'Noether currents'; the spatial components of these being $T^{i\mu}$, whose spatial divergence $T^{i\mu}{}_{,i}$ gives the rate of change of the corresponding density $T^{0\mu}$. Integrating the continuity equations over space gives the laws of conservation total energy and momentum

$$\int d^3x \, T^{\nu\mu}{}_{,\nu} = \frac{d}{dx^0} \int d^3x \, T^{0\mu} = 0. \tag{13}$$

One thing to note about the definition (10) is that it is not, in general, symmetric. We saw this in the case of electromagnetism, for which $\mathcal{L} = (16\pi\mu_0)^{-1} F^{\mu\nu} F_{\nu\mu}$ with $F_{\mu\nu} \equiv A_{\mu,\nu} - A_{\nu,\mu}$ the Faraday tensor. But we saw there that this could be converted into a symmetric – and gauge invariant – form by adding a term with vanishing 4-divergence. So there is, in general, some ambiguity about the stress-energy tensor.

2.2 Example: the relativistic scalar field

Let us consider, as an example of a matter field, a relativistic classical scalar field (for example the Higgs field or the axion), whose Lagrangian density in flat space, we may recall, is the Lorentz-scalar density

$$\mathcal{L}(\phi, \phi_{,\mu}) = -\tfrac{1}{2} \phi_{,\mu} \phi^{,\mu} - \tfrac{1}{2} m^2 \phi^2. \tag{14}$$

To remind ourselves of the physical meaning of the various terms encapsulated in this elegantly covariant equation, note that in 3+1 form, and letting dot denote derivative with respect to $x^0 = ct$, this is $\mathcal{L}(\phi, \dot{\phi}, \boldsymbol{\nabla}\phi) = \frac{1}{2}(\dot{\phi}^2 - |\boldsymbol{\nabla}\phi|^2 - m^2\phi^2)$. The space-integral of this is mathematically equivalent to the Lagrangian of a lattice of particles confined to their lattice locations with springs with potential energy $U_m = \sum_i m^2 \phi_i^2/2 \Rightarrow \int d^3x \, m^2\phi^2/2$ (so m^2 is the spring constant) and with connecting springs giving an additional potential $U_\Delta = \sum_i \Delta\phi_i^2/2 = \sum_i (\phi_{i+1} - \phi_i)^2/2 \Rightarrow \int d^3x \, |\boldsymbol{\nabla}\phi|^2/2$. With kinetic energy $K = \sum_i \dot{\phi}_i^2 m/2 \Rightarrow \int d^3x \, \dot{\phi}^2/2$ the Lagrangian for this 'scalar-elasticity' model is $L = K - (U_m + U_\Delta) = K - U = \int d^3x \mathcal{L}$ with $\mathcal{L}$ as given above.

Requiring that the variation of the action with respect to a variation of the field $\delta\phi(\vec{x})$ vanish gives the equation of motion for the field – or the Euler-Lagrange equation – here called the Klein-Gordon equation

$$\Box\phi = m^2\phi \tag{15}$$

and which allows wave-like solutions $\phi \sim e^{ik_\mu x^\mu} = e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_{\mathbf{k}} t)}$ provided $\omega_{\mathbf{k}} = ck_0$ obeys the dispersion relation $\omega_{\mathbf{k}}^2 = c^2(|\mathbf{k}|^2 + m^2)$.

The stress-energy tensor is

$$\begin{aligned} T^{\mu\nu} &\equiv -\phi^{,\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\nu}} + \eta^{\mu\nu} \mathcal{L} \\ &= \phi^{,\mu}\phi^{,\nu} + \eta^{\mu\nu} \left(-\tfrac{1}{2}\phi_{,\alpha}\phi^{,\alpha} - \tfrac{1}{2} m^2\phi^2 \right) \end{aligned} \tag{16}$$

or, in component form

$$T^{\mu\nu} = \begin{bmatrix} \frac{1}{2}(\dot{\phi}^2 + |\boldsymbol{\nabla}\phi|^2 + m^2\phi^2) & -\dot{\phi}\boldsymbol{\nabla}\phi \\ -\dot{\phi}\boldsymbol{\nabla}\phi & \boldsymbol{\nabla}\phi\boldsymbol{\nabla}\phi - \frac{1}{2}(\dot{\phi}^2 - |\boldsymbol{\nabla}\phi|^2 - m^2\phi^2)\mathbf{I} \end{bmatrix} \tag{17}$$

where $\mathbf{I}$ is the 3×3 identity matrix. So here the stress energy tensor is symmetric.

Just as the Lagrangian in the scalar elasticity model is $L = K - U = \int d^3x \frac{1}{2}(\dot{\phi}^2 - |\boldsymbol{\nabla}\phi|^2 - m^2\phi^2)$ the total energy is $H = K + U = \int d^3x \frac{1}{2}(\dot{\phi}^2 + |\boldsymbol{\nabla}\phi|^2 + m^2\phi^2)$, so T^{00} is the energy density and its integral over space is the conserved total energy H.

Of course we could have obtained this much more simply using the tools of ordinary classical mechanics where, from the Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, we define the generalised momentum $\mathbf{p} \equiv \partial L/\partial \dot{\mathbf{q}}$ and the Hamiltonian $H(\mathbf{q}, \mathbf{p}, t) \equiv \dot{\mathbf{q}} \cdot \mathbf{p} - L$. Then considering the change in the Hamiltonian $dH = (\partial H/\partial \mathbf{q}) \cdot d\mathbf{q} + (\partial H/\partial \mathbf{p}) \cdot d\mathbf{p} + (\partial H/\partial t) dt$ and invoking Hamilton's equations (equivalent to the Euler-Lagrange equation) $d\mathbf{q}/dt = \partial H/\partial \mathbf{p}$ and $d\mathbf{p}/dt = -\partial H/\partial \mathbf{x}$ we get $dH/dt = \partial H/\partial t = -\partial L/\partial t$. So if the Lagrangian we started with has no explicit dependence on time t, the Hamiltonian is conserved: $dH/dt = 0$.

In the scalar-elasticity model (and considering one spatial dimension for simplicity) the components of the vector $\mathbf{q}$ are the values of the field at the lattice points ϕ_i, the components of $\mathbf{p}$ are the $\dot{\phi}_i$, and a dot product like $\mathbf{q} \cdot \mathbf{p} = \sum_i q_i p_i$, or a term in the Lagrangian like the kinetic energy $T = \dot{\mathbf{q}} \cdot \dot{\mathbf{q}}/2 = \sum_i \dot{q}_i^2/2$, is a sum over lattice points or, in the 'continuum limit', an integral over space. The Lagrangian $L = K - U = \sum_i (\dot{\phi}_i^2 - \Delta\phi_i^2 - m^2\phi_i^2)/2$ is then $L = \sum_i \mathcal{L}_i \Rightarrow \int dx \mathcal{L}$ and similarly the Hamiltonian is $H = \sum_i \mathcal{H}_i \Rightarrow \int dx \mathcal{H}$ where $\mathcal{H}_i = (\dot{\phi}_i^2 + \Delta\phi_i^2 + m^2\phi_i^2)/2$ is the Hamiltonian density.

What the standard classical mechanics does not yield quite so easily is the form and physical significance of the other terms in the stress-energy tensor. The vanishing of the 4-divergence of the components of the 0th column ${T^{\nu 0}}_{,\nu}$ is saying $\partial_0 T^{00} + \boldsymbol{\nabla} \cdot (-\dot{\phi} \boldsymbol{\nabla} \phi) = 0$ or that the rate of change of energy density is $\dot{\mathcal{H}} = -\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{F}}$ where the energy flux density is $\boldsymbol{\mathcal{F}} \equiv -\dot{\phi} \boldsymbol{\nabla} \phi$ (which, in electromagnetism, is called the 'Poynting-flux'). Similarly, the vanishing 4-divergence of the components of the spatial columns ${T^{\nu i}}_{,\nu} = 0$ are expressed as $\dot{\boldsymbol{\mathcal{P}}} = -\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{S}}$ where the momentum density is defined as $\boldsymbol{\mathcal{P}} \equiv -\dot{\phi} \boldsymbol{\nabla} \phi$ (which is the same as the energy flux density $\boldsymbol{\mathcal{F}}$ for a scalar field) and the momentum flux density – the 3-stress tensor – has components $\mathcal{S}_{ij} = T_{ij}$.

If we think about a nearly-monochromatic wave-packet varying locally as $\phi \sim e^{ik_\mu x^\mu}$ as we then find its energy is $H = \omega_{\mathbf{k}} \times \omega_{\mathbf{k}} \int d^3r \phi^2$ while its total 3-momentum is $\mathbf{p} = \mathbf{k} \times \omega_{\mathbf{k}} \int d^3r \phi^2$. From this one can show that such packets have total energy and momentum that obey the relativistic relation $H^2 = p^2 + m^2$.

3 The gravitational action

The Ricci scalar is $R = g^{\mu\nu} R_{\mu\nu}$ where $R_{\mu\nu} = {R^\alpha}_{\mu\alpha\nu}$ with ${R^\alpha}_{\mu\beta\nu} = {\Gamma^\alpha}_{\mu\nu,\beta} - {\Gamma^\alpha}_{\mu\beta,\nu} + {\Gamma^\alpha}_{\gamma\beta}{\Gamma^\gamma}_{\mu\nu} - {\Gamma^\alpha}_{\gamma\nu}{\Gamma^\gamma}_{\mu\beta}$ and where ${\Gamma^\alpha}_{\mu\nu} = \frac{1}{2} g^{\alpha\beta}(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta})$.

In mechanics, the action is $S = \int dt L$, where Lagrangian L is a function of the generalised coordinates $\mathbf{q}$, their velocities $\dot{\mathbf{q}}$, and time: $L = L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The velocities usually enter quadratically, and this results in equations of motion that are second order in time. Similarly, for fields, the Lagrangian *density* – the integrand in $S = \int dt \int d^3x \mathcal{L}(\phi, \phi_{,\mu})$ – contains the fields and, again quadratically, their first derivatives with respect to space-time coordinates. Here the metric will play the role of the generalised coordinates or the fields and the Ricci scalar that of the Lagrangian density. But while the Ricci scalar contains products of derivatives of the metric components in the terms coming from products of Christoffel symbols, it also contains second derivatives via the first two terms in the curvature tensor, which seems problematic. But these only appear at first order, and, as we now show, the Ricci scalar can be written as a combination of terms that are quadratic in the metric first derivatives plus total derivatives. These give only 'boundary terms' in the action integral, so, if we consider the fields to be held fixed on the boundary, or imagine the boundary either to be at infinity – where, as usual, we will consider the fields to vanish – or, as in the case of a closed cosmology, not to exist at all, then these do not contribute to the variation of the action.

Following Dirac, we split gravitational action into two parts:

$$S_{\rm g} \equiv \frac{1}{16\pi\kappa} \int d^4x \sqrt{g} R = \int d^4x (L^\star - L) \tag{18}$$

where

$$\begin{aligned} L^\star &\equiv \frac{1}{16\pi\kappa} \sqrt{g} g^{\mu\nu} ({\Gamma^\alpha}_{\mu\nu,\alpha} - {\Gamma^\alpha}_{\mu\alpha,\nu}) \\ L &\equiv \frac{1}{16\pi\kappa} \sqrt{g} g^{\mu\nu} ({\Gamma^\alpha}_{\gamma\nu}{\Gamma^\gamma}_{\mu\alpha} - {\Gamma^\alpha}_{\gamma\alpha}{\Gamma^\gamma}_{\mu\nu}) \end{aligned} \tag{19}$$

The first of these – which contains the unwanted second derivatives of the metric – can be written as

$$L^\star = \cancel{(\sqrt{g} g^{\mu\nu} {\Gamma^\alpha}_{\mu\nu})_{,\alpha}} - \cancel{(\sqrt{g} g^{\mu\nu} {\Gamma^\alpha}_{\mu\alpha})_{,\nu}} - {\Gamma^\alpha}_{\mu\nu} (\sqrt{g} g^{\mu\nu})_{,\alpha} + {\Gamma^\alpha}_{\mu\alpha} (\sqrt{g} g^{\mu\nu})_{,\nu} \tag{20}$$

where the first pair of terms are the total derivates, which we have struck out, not because they cancel each other, but to indicate that they do not contribute to δS.

We would like to express the remaining two terms as products of Christoffel symbols. To do this, we will use the fact that the metric is covariantly constant so $g^{\mu\nu}{}_{;\alpha} = 0 \Rightarrow g^{\mu\nu}{}_{,\alpha} = -\Gamma^\mu{}_{\gamma\alpha} g^{\gamma\nu} - \Gamma^\nu{}_{\gamma\alpha} g^{\mu\gamma}$ and the fact that $(\sqrt{g})_{,\alpha} = \frac{1}{2}\sqrt{g} g^{\gamma\delta} g_{\gamma\delta,\alpha} = \sqrt{g}\Gamma^\nu{}_{\alpha\nu}$, which follows from linear algebra as you showed in a homework problem. These give

$$\begin{aligned}(\sqrt{g}g^{\mu\nu})_{,\alpha} &= \sqrt{g}g^{\mu\nu}{}_{,\alpha} + g^{\mu\nu}(\sqrt{g})_{,\alpha} \\ &= \sqrt{g}(-\Gamma^\mu{}_{\gamma\alpha}g^{\gamma\nu} - \Gamma^\nu{}_{\gamma\alpha}g^{\mu\gamma} + \tfrac{1}{2}\Gamma^\gamma{}_{\alpha\gamma}g^{\mu\nu})\end{aligned} \tag{21}$$

which, in (20), gives 6 terms, two of which cancel, leaving 4 which are two duplicate pairs, that turn out to be $2L$. The result is that the variation of the gravitational action is

$$\delta S_{\rm g} = \frac{1}{16\pi\kappa}\delta\int d^4x\sqrt{g}R = \delta\int d^4x L(g_{\rho\sigma}, g_{\rho\sigma,\tau}) \tag{22}$$

More often, you will see $\int d^4xL$ is written as $\int d^4x\sqrt{g}\mathcal{L}_{\rm g}$, where $\mathcal{L}_{\rm g} \equiv g^{\mu\nu}(\Gamma^\alpha{}_{\gamma\nu}\Gamma^\gamma{}_{\mu\alpha} - \Gamma^\alpha{}_{\gamma\alpha}\Gamma^\gamma{}_{\mu\nu})/16\pi\kappa$, and which, from the definition of the Christoffel symbols, is a function of $g^{\rho\sigma}$ and $g_{\rho\sigma,\tau}$. However, since the inverse metric and the determinant are both functions of $g_{\rho\sigma}$, we can consider L, as indicated, to be a function of $g_{\rho\sigma}$ and $g_{\rho\sigma,\tau}$.

We can now proceed exactly as we did for the matter fields. The variation of $L_{\rm g}$ is simply

$$\begin{aligned}\delta L &= \frac{\partial L}{\partial g_{\rho\sigma,\tau}}\delta g_{\rho\sigma,\tau} + \frac{\partial L}{\partial g_{\rho\sigma}}\delta g_{\rho\sigma} \\ &= \cancel{\left(\frac{\partial L}{\partial g_{\rho\sigma,\tau}}\delta g_{\rho\sigma}\right)_{,\tau}} - \left(\frac{\partial L}{\partial g_{\rho\sigma,\tau}}\right)_{,\tau}\delta g_{\rho\sigma} + \frac{\partial L}{\partial g_{\rho\sigma}}\delta g_{\rho\sigma}\end{aligned} \tag{23}$$

since $\partial_\gamma\delta g_{\rho\sigma} = \delta g_{\rho\sigma,\tau}$ and where we have 'differentiated by parts' to eliminate $\delta g_{\rho\sigma,\tau}$. It is still there, of course, but inside the total derivate term, which, just like the total derivatives in $L^\star$, do not contribute to $\delta S_{\rm g}$, which is why it is struck out. Thus the variation of the gravitational action is

$$\delta S_{\rm g} = \int d^4x\delta g_{\rho\sigma}\left(\frac{\partial L}{\partial g_{\rho\sigma}} - \partial_\tau\frac{\partial L}{\partial g_{\rho\sigma,\tau}}\right) \tag{24}$$

and requiring that $\delta S_{\rm g}$ vanish for an arbitrary variation $\delta g_{\rho\sigma}$ gives Einstein's equations in a vacuum.

We can evaluate the quantity in parentheses – let's call it $-\tilde{G}^{\rho\sigma}/16\pi\kappa$ – most simply in a locally inertial frame (where the Christoffel symbols vanish and $g^{\rho\sigma} = \eta^{\rho\sigma}$ and $\sqrt{g} = 1$). Because L contains products of Christoffel symbols, $\partial L/\partial g_{\rho\sigma} = 0$, and we can similarly ignore the derivatives of $\sqrt{g}g^{\mu\nu}$ in the other term as these would get multiplied by products of Christoffel symbols. The only non-vanishing terms are those containing second derivatives of the metric – and we have

$$\tilde{G}^{\rho\sigma} = \eta^{\mu\nu}\partial_\gamma\frac{\partial}{\partial g_{\rho\sigma,\tau}}\left(\tilde{\Gamma}^\alpha_{\gamma\nu}\tilde{\Gamma}^\gamma_{\mu\alpha} - \tilde{\Gamma}^\alpha_{\gamma\alpha}\tilde{\Gamma}^\gamma_{\mu\nu}\right) \tag{25}$$

where the $\tilde{\Gamma}^\alpha_{\gamma\nu} \equiv \frac{1}{2}\eta^{\alpha\beta}(g_{\beta\gamma,\nu} + g_{\beta\nu,\gamma} - g_{\gamma\nu,\beta})$ are the Chrisfoffel symbols with $g^{\alpha\beta}$ replaced by $\eta^{\alpha\beta}$.

It is now a little tedious, but conceptually straightforward, to perform the derivatives here. The first term is

$$\begin{aligned}&\partial_\tau\eta^{\mu\nu}\partial(\tilde{\Gamma}^\alpha_{\mu\gamma}\tilde{\Gamma}^\gamma_{\nu\alpha})/\partial g_{\rho\sigma,\tau} = \partial_\tau\eta^{\mu\nu}\eta^{\alpha\beta}\eta^{\gamma\pi}\partial(\tilde{\Gamma}_{\beta\mu\gamma}\tilde{\Gamma}_{\pi\nu\alpha})/\partial g_{\rho\sigma,\tau} \\ &\quad= \tfrac{1}{2}\partial_\tau\eta^{\mu\nu}\eta^{\alpha\beta}\eta^{\gamma\pi}[\tilde{\Gamma}_{\beta\mu\gamma}(\delta^\rho_\pi\delta^\sigma_\nu\delta^\tau_\alpha + \delta^\rho_\pi\delta^\sigma_\alpha\delta^\tau_\nu - \delta^\rho_\alpha\delta^\sigma_\nu\delta^\tau_\pi) + \tilde{\Gamma}_{\pi\nu\alpha}(\delta^\rho_\beta\delta^\sigma_\mu\delta^\tau_\gamma + \delta^\rho_\beta\delta^\sigma_\gamma\delta^\tau_\mu - \delta^\rho_\gamma\delta^\sigma_\mu\delta^\tau_\beta)] \\ &\quad= \partial_\tau(\tilde{\Gamma}^{\tau\sigma\rho} + \tilde{\Gamma}^{\sigma\tau\rho} - \tilde{\Gamma}^{\rho\sigma\tau}) = \tfrac{1}{2}(3g^{\sigma\tau,\rho}{}_{,\tau} - g^{\rho\tau,\sigma}{}_{,\tau} - g^{\rho\sigma,\tau}{}_{,\tau})\end{aligned} \tag{26}$$

while (minus) the second is

$$\begin{aligned}&\partial_\tau\eta^{\mu\nu}\frac{\partial(\tilde{\Gamma}^\alpha_{\gamma\alpha}\tilde{\Gamma}^\gamma_{\mu\nu})}{\partial g_{\rho\sigma,\tau}} = \tfrac{1}{2}\partial_\tau\eta^{\mu\nu}\eta^{\alpha\beta}\eta^{\gamma\pi}\frac{\partial(g_{\alpha\beta,\gamma}\tilde{\Gamma}_{\pi\mu\nu})}{\partial g_{\rho\sigma,\tau}} = \tfrac{1}{4}\partial_\tau\eta^{\mu\nu}\eta^{\alpha\beta}\eta^{\gamma\pi}\frac{\partial(g_{\alpha\beta,\gamma}(2g_{\pi\mu,\nu} - g_{\mu\nu,\pi}))}{\partial g_{\rho\sigma,\tau}} \\ &\quad= \tfrac{1}{4}\partial_\tau\eta^{\mu\nu}\eta^{\alpha\beta}\eta^{\gamma\pi}[\delta^\rho_\alpha\delta^\sigma_\beta\delta^\tau_\gamma(2g_{\pi\mu,\nu} - g_{\mu\nu,\pi}) + (2\delta^\rho_\pi\delta^\sigma_\mu\delta^\tau_\nu - \delta_\mu\rho\delta^\sigma_\nu\delta^\tau_\pi)g_{\alpha\beta,\gamma}] \\ &\quad= \tfrac{1}{2}[g^{,\rho\sigma} + \eta^{\rho\sigma}(g^{\mu\tau}{}_{,\mu\tau} - g^{,\tau}{}_{,\tau})]\end{aligned} \tag{27}$$

subtracting the second from the first gives

$$\begin{aligned}\tilde{G}^{\rho\sigma} &= -\tfrac{1}{2}[g^{,\rho\sigma} - g^{\rho\tau,\sigma}{}_{,\tau} + g^{\rho\sigma,\tau}{}_{,\tau} - g^{\sigma\tau,\rho}{}_{,\tau}] - \tfrac{1}{2}\eta^{\rho\sigma}(g^{\mu\tau}{}_{,\mu\tau} - g^{,\tau}{}_{,\tau}) + (g^{\rho\tau,\sigma}{}_{,\tau} - g^{\sigma\tau,\rho}{}_{,\tau}) \\ &= G^{\rho\sigma} + (g^{\rho\tau,\sigma}{}_{,\tau} - g^{\sigma\tau,\rho}{}_{,\tau})\end{aligned} \tag{28}$$

So this isn't identical to the Einstein tensor – it contains an additional asymmetric term – but recall that what we have calculated here is the quantity in parentheses in (24). That appears contracted with the symmetric $\delta g_{\rho\sigma}$ so the asymmetric terms cancel, and we obtain

$$\boxed{\delta S_{\rm g} = -\frac{1}{16\pi\kappa}\int d^4x\sqrt{g}G^{\mu\nu}\delta g_{\mu\nu}} \tag{29}$$

where we have used the fact that $d^4x = \sqrt{g}d^4x$ in locally inertial coordinates, and which is what we set out to prove.

4 The stress-energy tensor as the derivative of $\mathcal{L}_{\rm m}$ with respect to the metric

We have seen that the variation of the gravitational part of the Einstein-Hilbert action (1) is

$$\delta(\sqrt{g}\mathcal{L}_{\rm grav}) = \frac{1}{16\pi\kappa}\delta(\sqrt{g}R) = -\frac{1}{16\pi\kappa}\sqrt{g}G^{\mu\nu}\delta g_{\mu\nu}. \tag{30}$$

The action principle says

$$\delta(\sqrt{g}(\mathcal{L}_{\rm grav} + \mathcal{L}_{\rm m})) = 0. \tag{31}$$

For this to be correct, and compatible with Einstein's equations $G^{\mu\nu} = 8\pi\kappa T^{\mu\nu}$, it must be that

$$\delta(\sqrt{g}\mathcal{L}_{\rm m}) = \tfrac{1}{2}\sqrt{g}T^{\mu\nu}\delta g_{\mu\nu} \tag{32}$$

or that

$$T^{\mu\nu} = \frac{2}{\sqrt{g}}\frac{\delta(\sqrt{g}\mathcal{L}_{\rm m})}{\delta g_{\mu\nu}} \tag{33}$$

or

$$T^{\mu\nu} = 2\frac{\delta\mathcal{L}_{\rm m}}{\delta g_{\mu\nu}} + g^{\mu\nu}\mathcal{L}_{\rm m} \tag{34}$$

where we have used $\delta\sqrt{g} = \frac{1}{2}\sqrt{g}g^{\nu\sigma}\delta g_{\nu\sigma}$.

Comparing with the formula for the stress-energy tensor in flat space time (or, equivalently, in a locally inertial frame, where we have $g^{\mu\nu} = \eta^{\mu\nu}$) $T^{\mu\nu} = -\phi^{,\mu}\partial\mathcal{L}_{\rm m}/\partial\phi_{,\nu} + \eta^{\mu\nu}\mathcal{L}_{\rm m}$ it must be that $2\delta\mathcal{L}_{\rm m}/\delta g_{\mu\nu}$ and $-\phi^{,\mu}\partial\mathcal{L}_{\rm m}/\partial\phi_{,\nu}$ are the same.

But that is correct, as we can see by considering e.g. the kinetic term $\mathcal{L}_{\rm kin} = -\frac{1}{2}\phi_{,\mu}\phi^{,\mu} = -\frac{1}{2}g^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$ for the scalar field. The variation of this with respect to a variation of the metric, holding $\phi_{,\mu}$ fixed is $\delta\mathcal{L}_{\rm kin} = -\frac{1}{2}\phi_{,\mu}\phi_{,\nu}\delta g^{\mu\nu}$ which, with $\delta g^{\nu\mu} = -g^{\nu\alpha}g^{\mu\beta}\delta g_{\alpha\beta}$, is $\delta\mathcal{L}_{\rm kin} = \frac{1}{2}\phi^{,\mu}\phi^{,\nu}\delta g_{\mu\nu}$ and hence $2\delta\mathcal{L}_{\rm kin}/\delta g_{\mu\nu} = \phi^{,\mu}\phi^{,\nu}$ while $-\phi^{,\mu}\partial\mathcal{L}_{\rm kin}/\partial\phi_{,\nu} = \phi^{,\mu}\phi^{,\nu}$ also.