A scalar field in curved spacetime

Joachim Kopp*

Max Planck Institut für Kernphysik,
Saupfercheckweg 1, 69117 Heidelberg, Germany
(Dated: May 3, 2013)

In this supplementary note to the Theoretical Astroparticle Physics lecture from May 02, 2013, we clarify a few points regarding the derivation of Einstein's equations from a Lagrangian, and regarding the energy-momentum tensor of a scalar field.

1. EINSTEIN'S EQUATION FROM A LAGRANGIAN

In the following, we derive Einstein's equations from the action

$$S = \int d^4x \, (\mathcal{L}_{\rm GR} + \mathcal{L}_{\rm matter}) \,, \tag{1}$$

where

$$\mathcal{L}_{\rm GR} = -\frac{1}{16\pi G} \sqrt{-g} (R + 2\Lambda) \,, \tag{2}$$

$$\mathcal{L}_{\rm matter} = \sqrt{-g} \, \tilde{\mathcal{L}} = \sqrt{-g} \left[(\partial_\mu \phi^\dagger)(\partial_\nu \phi) g^{\mu\nu} - V(\phi^\dagger \phi) \right] . \tag{3}$$

Here, G is the gravitational constant, R is the Ricci scalar, Λ is the cosmological constant, ϕ is a complex scalar field with potential $V(\phi^\dagger\phi)$, and g is the determinant of the metric. $\tilde{\mathcal{L}}$ is the matter Lagrangian in Minkowski space. Note that for a scalar field, the covariant derivative is identical to the ordinary derivative, hence we write ∂_μ instead of ∇_μ for simplicity. (When taking derivatives of vector fields, one has to use the covariant derivative ∇_μ in order to ensure diffeomorphism invariance of the Lagrangian.)

The variation of $\mathcal{L}_{\rm GR}$ is

$$\delta\mathcal{L}_{\rm GR} = -\frac{1}{16\pi G} \left[\delta(\sqrt{-g})(R + 2\Lambda) + \sqrt{-g} \, \delta(g_{\mu\nu} R^{\mu\nu}) \right] . \tag{4}$$

To evaluate the first term, we use

$$\delta(\sqrt{-g}) = \frac{1}{2} \sqrt{-g} \, g^{\mu\nu} \delta g_{\mu\nu} \,, \tag{5}$$

which follows from the matrix identity

$$\frac{\partial \det A(x)}{\partial x} = \det A(x) \times {\rm tr} \left[A^{-1}(x) \, \frac{\partial A(x)}{\partial x} \right] \tag{6}$$

for an arbitrary matrix A depending on a parameter x.

It can be shown that the second term in eq. (4) yields

$$-\sqrt{-g} R^{\mu\nu} \delta g_{\mu\nu} \,, \tag{7}$$

*Email: jkopp@mpi-hd.mpg.de

plus a total covariant derivative which can be omitted from the action if we apply Stokes' theorem and assume as usual that boundary terms vanish.

To see why the second term in eq. (4) has this simple form, one has to begin with the definition of the Riemann curvature tensor in terms of the Christoffel symbols,

$$R_{\mu\nu\rho}{}^{\sigma} = \Gamma^{\sigma}{}_{\mu\rho,\nu} - \Gamma^{\sigma}{}_{\nu\rho,\mu} + \Gamma^{\alpha}{}_{\mu\rho}\Gamma^{\sigma}{}_{\alpha\nu} - \Gamma^{\alpha}{}_{\nu\rho}\Gamma^{\sigma}{}_{\alpha\mu} \,. \tag{8}$$

The variation of $R_{\mu\nu\rho}{}^{\sigma}$ is

$$\delta R_{\mu\nu\rho}{}^{\sigma} = \partial_\nu \delta\Gamma^{\sigma}{}_{\mu\rho} - \partial_\mu \delta\Gamma^{\sigma}{}_{\nu\rho} + (\delta\Gamma^{\alpha}{}_{\mu\rho})\Gamma^{\sigma}{}_{\alpha\nu} + \Gamma^{\alpha}{}_{\mu\rho}(\delta\Gamma^{\sigma}{}_{\alpha\nu}) - (\delta\Gamma^{\alpha}{}_{\nu\rho})\Gamma^{\sigma}{}_{\alpha\mu} - \Gamma^{\alpha}{}_{\nu\rho}(\delta\Gamma^{\sigma}{}_{\alpha\mu}) \,. \tag{9}$$

We now use the definition of the covariant derivative of an arbitrary tensor $t^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\beta_n}$,

$$\nabla_\nu t^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\beta_n} = \partial_\nu t^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\beta_n} + \sum_i \Gamma^{\alpha_i}{}_{\nu\tau} t^{\alpha_1\cdots\tau\cdots\alpha_n}{}_{\beta_1\cdots\beta_n} - \sum_i \Gamma^{\tau}{}_{\nu\beta_i} t^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\tau\cdots\beta_n} \tag{10}$$

to show that

$$\nabla_\nu \, \delta\Gamma^{\sigma}{}_{\mu\rho} = \partial_\nu \delta\Gamma^{\sigma}{}_{\mu\rho} + \Gamma^{\sigma}{}_{\nu\alpha}\, \delta\Gamma^{\alpha}{}_{\mu\rho} - \Gamma^{\alpha}{}_{\nu\mu}\, \delta\Gamma^{\sigma}{}_{\alpha\rho} - \Gamma^{\alpha}{}_{\nu\rho}\, \delta\Gamma^{\sigma}{}_{\alpha\mu} \tag{11}$$

and thus eq. (9) reduces to

$$\delta R_{\mu\nu\rho}{}^{\sigma} = \nabla_\nu \, \delta\Gamma^{\sigma}{}_{\mu\rho} - \nabla_\mu \, \delta\Gamma^{\sigma}{}_{\nu\rho} \,. \tag{12}$$

Here, we have used the identity $\Gamma^{\alpha}{}_{\beta\gamma} = \Gamma^{\alpha}{}_{\gamma\beta}$ as well as the fact that, while a Christoffel symbol $\Gamma^{\alpha}{}_{\beta\gamma}$ is not a Lorentz tensor, its variation $\delta\Gamma^{\alpha}{}_{\beta\gamma}$ is. (Hence we can form its covariant derivative.) To see this, write $\delta(\Gamma^{\sigma}{}_{\mu\rho} t^{\rho})$ for an arbitrary vector field t^{ρ} in terms of partial and covariant derivatives:

$$(\delta\Gamma^{\sigma}{}_{\mu\rho})t^{\rho} + \Gamma^{\sigma}{}_{\mu\rho}\delta t^{\rho} = \delta(\nabla_\mu t^{\sigma}) - \partial_\mu \delta t^{\rho} \,, \tag{13}$$

from which it follows that

$$(\delta\Gamma^{\sigma}{}_{\mu\rho})t^{\rho} = \delta(\nabla_\mu t^{\sigma}) - \nabla_\mu \delta t^{\rho} \,. \tag{14}$$

The right hand side of this equation is manifestly covariant, so the left hand side must be covariant as well. Eq. (12) immediately yields the variation of the Ricci tensor $R_{\mu\rho} = R_{\mu\alpha\rho}{}^{\alpha}$:

$$\delta R_{\mu\rho} = \nabla_\alpha \, \delta\Gamma^{\alpha}{}_{\mu\rho} - \nabla_\mu \, \delta\Gamma^{\alpha}{}_{\alpha\rho} \tag{15}$$

and of the Ricci scalar $R = R^{\mu}{}_{\mu}$:

$$\delta R = \delta(R_{\mu\rho} g^{\mu\rho}) = (\nabla_\alpha \, \delta\Gamma^{\alpha}{}_{\mu\rho}) g^{\mu\rho} - (\nabla_\mu \, \delta\Gamma^{\alpha}{}_{\alpha\rho}) g^{\mu\rho} - R^{\mu\rho}\delta g_{\mu\rho} \,. \tag{16}$$

In the last term, we have used $\delta g^{\mu\rho} = -g^{\mu\alpha}\, g^{\nu\beta}\, \delta g_{\alpha\beta}$, which in turn follows from the well-known relation for the derivative of the inverse of a matrix: $\partial A^{-1}(x)/\partial(x) = -A^{-1}(\partial A/\partial x)A^{-1}$. The first two terms on the right hand side of eq. (16) form a total covariant derivative because $g^{\mu\rho}$ can be pulled under the covariant derivative since $\nabla_\alpha g^{\mu\rho} = 0$. By the same argument, these terms still yield a total covariant derivative when multiplied by $\sqrt{-g}$, as in eq. (4).

Plugging everything into eq. (4) and neglecting the total derivative terms we obtain

$$\delta\mathcal{L}_{\mathrm{GR}} = -\frac{1}{16\pi G}\left[\frac{1}{2}(R + 2\Lambda)g^{\mu\nu} - R^{\mu\nu}\right]\sqrt{-g}\,\delta g_{\mu\nu} \,. \tag{17}$$

The term in square brackets is just the Einstein tensor—the left hand side of Einstein's equation. To obtain the right hand side, we now vary $\mathcal{L}_{\mathrm{matter}}$ with respect to $g_{\mu\nu}$:

$$\delta\mathcal{L}_{\mathrm{matter}} = \delta(\sqrt{-g}\tilde{\mathcal{L}}) \tag{18}$$

$$= \frac{1}{2}\sqrt{-g}\, g^{\mu\nu}\tilde{\mathcal{L}}\,\delta g_{\mu\nu} + \frac{\delta\tilde{\mathcal{L}}}{\delta g_{\mu\nu}}\sqrt{-g}\,\delta g_{\mu\nu}\,. \tag{19}$$

Adding up eqs. (19) and (17), and using that the variation of the action has to vanish for any $\delta g_{\mu\nu}$, we obtain Einstein's equation:

$$R^{\mu\nu} - \frac{R}{2}g^{\mu\nu} = 16\pi G\Big[-\frac{1}{2}g^{\mu\nu}\tilde{\mathcal{L}} - \frac{\delta\tilde{\mathcal{L}}}{\delta g_{\mu\nu}} + \frac{\Lambda}{16\pi G}\Big] \tag{20}$$

$$= 8\pi G T^{\mu\nu} + \Lambda g^{\mu\nu}\,. \tag{21}$$

In the last step, we have defined the stress-energy tensor

$$T^{\mu\nu} = -g^{\mu\nu}\tilde{\mathcal{L}} - 2\frac{\delta\tilde{\mathcal{L}}}{\delta g_{\mu\nu}}\,. \tag{22}$$

2. THE STRESS-ENERGY TENSOR FOR A SCALAR FIELD

We now apply eq. (22) to the case of a scalar field with a Lagrangian given by eq. (3). We obtain

$$T^{\mu\nu} = -(\partial_\rho\phi^\dagger)(\partial^\rho\phi)g^{\mu\nu} + V(\phi^\dagger\phi)g^{\mu\nu} + 2(\partial_\alpha\phi^\dagger)(\partial_\beta\phi)g^{\alpha\mu}g^{\beta\nu}\,. \tag{23}$$

Note that in the last term, we have again used $\delta g^{\mu\rho} = -g^{\mu\alpha}\, g^{\nu\beta}\,\delta g_{\alpha\beta}$.

It is important here that we write the kinetic term as $(\partial_\alpha\phi^\dagger)(\partial_\beta\phi)g^{\alpha\beta}$ before varying it with respect to $g_{\mu\nu}$, not as $(\partial^\alpha\phi^\dagger)(\partial^\beta\phi)g_{\alpha\beta}$. The reason is that a Lagrangian is always a function of fields (ϕ in our example) and their first covariant (not cotravariant) derivatives with respect to the space-time coordinates ($\nabla_\mu\phi = \partial_\mu\phi$ in our case). If we chose to treat ϕ and the contravariant derivative $\nabla^\mu\phi = \partial^\mu\phi$ as the fundamental functions instead, we would get into trouble when attempting to derive the Euler-Lagrange equations for the field. Doing so requires applying Stokes' theorem once to remove a total derivative term. Stokes' theorem, however, holds only for total covariant derivatives (= derivatives with *lower* indices).

In a Friedmann-Robertson-Walker Universe, we obtain for the time component T^{00}:

$$T^{00} = (\partial_0\phi)^\dagger(\partial_0\phi) + \frac{1}{R^2(t)}\sum_j(\partial_j\phi)^\dagger(\partial_j\phi) + V(\phi^\dagger\phi)\,, \tag{24}$$

and for the spatial components T^{ii}:

$$T^{ii} = \frac{1}{R^2(t)}(\partial_0\phi)^\dagger(\partial_0\phi) - \frac{1}{R^4(t)}\sum_j(\partial_j\phi)^\dagger(\partial_j\phi) + 2\frac{1}{R^4(t)}(\partial_i\phi)^\dagger(\partial_i\phi) - \frac{1}{R^2(t)}V(\phi^\dagger\phi)\,. \tag{25}$$

Note that spatial indices j are only summed where explicitly indicated. In the case of a perfect, homogeneous, isotropic fluid, the spatial derivate terms vanish. To interpret the components of T in terms of the energy density and the pressure in this case, we note that $(T^\mu{}_\nu) = \mathrm{diag}(\rho, -p, -p, -p)$ for a perfect homogeneous, isotropic fluid (note the position of the indices!). This leads to

$$\rho = |\dot\phi|^2 + V\,, \tag{26}$$

$$p = |\dot\phi|^2 - V\,. \tag{27}$$

3. ANSWERS TO QUESTIONS THAT AROSE DURING THE LECTURE ON 02.05.2013

1. **A missing minus sign in the derivation of $\delta\mathcal{L}_{\mathbf{GR}}$?**
The origin of the confusion was the way in which I had expanded $\delta(g_{\mu\nu}R^{\mu\nu})$ in eq. (4). It is important to use

$$\delta(g^{\mu\nu}R_{\mu\nu}) = R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} \tag{28}$$

and *not*

$$\delta(g_{\mu\nu}R^{\mu\nu}) = R^{\mu\nu}\delta g_{\mu\nu} + g_{\mu\nu}\delta R^{\mu\nu} \tag{29}$$

The reason is that $\delta R_{\mu\nu}$ yields a total covariant derivative ∇_α (with *lower* index), whereas $\delta R^{\mu\nu}$ yields a total contravariant derivative ∇^α (with *upper* index). In order to remove the total derivative from the action, we need to apply Stokes' theorem which holds for derivatives with *lower* indices.

Since we want to pull $\delta g_{\mu\nu}$ (with lower indices) out of the square brackets in eq. (17), we need to use $\delta g^{\mu\rho} = -g^{\mu\alpha}\, g^{\nu\beta}\, \delta g_{\alpha\beta}$ in the first term on the right hand side of eq. (28), hence the relative minus sign in eq. (17).

2. **Do we assume Minkowski space in deriving expressions for ρ and p?**
In the lecture, we assumed the Minkowski metric, which is always possible at a given fixed time t_1 in cosmological history if we rescale coordinates in such a way that $R(t_1) = 1$. However, this assumption is not necessary, as the derivation in sec. 2 shows. The extra factors of $R(t)$ that appear in eqs. (24) and (25) vanish when the spatial derivatives are discarded and when one of the indices is lowered to obtain $T^{\mu}{}_{\nu}$.