GR lecture 4-2

Free-falling particle, conservation laws, some cosmology, exterior & Lie derivatives

I. PARTICLE MOTION IN A CURVED METRIC

The action of a free particle in a curved metric is the same as its special-relativistic version: it is $-m$ times the length of the worldline. However, we should now make explicit the dependence of length on the metric:

$$S = -m \int d\tau = -m \int \sqrt{-g_{\mu\nu}(x) dx^\mu dx^\nu} \ . \tag{1}$$

Note that even a "free" particle has no choice but to feel the effects of spacetime geometry, i.e. of the gravitational field! First of all, let's examine this action in the limit of slow, non-relativistic motion, and weak gravitational fields. In other words, let's expand eq. (1) to leading order in $g_{\mu\nu} - \eta_{\mu\nu}$ and $\mathbf{v}$. Denoting $g_{tt}(t,\mathbf{x}) = -1 - 2\varphi(t,\mathbf{x})$, we have:

$$S \approx -m \int \sqrt{-g_{tt} dt^2 - \mathbf{dx}^2} = -m \int dt \sqrt{1 + 2\varphi - \mathbf{v}^2} \approx -m \int dt \left(1 + \varphi - \frac{\mathbf{v}^2}{2}\right) \ . \tag{2}$$

Thus, the Lagrangian for a non-relativistic particle in a weak gravitational field is:

$$L = -m + \frac{m\mathbf{v}^2}{2} - m\varphi(t,\mathbf{x}) \ . \tag{3}$$

The first term is a constant, and doesn't affect the equations of motion. The second is kinetic energy. The third has the form of potential energy. We conclude that the deviation $\varphi = -(g_{tt}+1)/2$ of g_{tt} from its flat value $\eta_{tt} = -1$ is the Newtonian gravitational potential! Note the analogy with how the non-relativistic electrostatic potential is actually the time component $-A_t$ of the electromagnetic potential A_μ.

Let's now obtain the full equation of motion from the action (1), without assuming slow motion or weak fields. We again introduce a parameter λ along the worldline, which can now be written as $x^\mu(\lambda)$, and denote λ derivatives as $\dot{x}^\mu \equiv dx^\mu/d\lambda$. The particle action then becomes:

$$S = \int L(x^\mu, \dot{x}^\mu)\, d\lambda \ ; \quad L = -m\sqrt{-g_{\mu\nu}(x)\dot{x}^\mu \dot{x}^\nu} \ . \tag{4}$$

The non-trivial metric introduces a dependence on x into the particle Lagrangian! The Euler-Lagrange equations of motion now read:

$$0 = \frac{\partial L}{\partial x^\mu} - \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = m\left(\frac{\partial_\mu g_{\nu\lambda}\dot{x}^\nu \dot{x}^\lambda}{2\sqrt{-g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}} - \frac{d}{d\lambda}\frac{g_{\mu\nu}\dot{x}^\nu}{\sqrt{-g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}}\right) . \tag{5}$$

Choosing λ to be the proper time τ, this becomes:

$$m\frac{du_\mu}{d\tau} = \frac{m}{2}(\partial_\mu g_{\nu\rho})u^\nu u^\rho . \tag{6}$$

On the LHS, we recognize $p_\mu = mu_\mu$ as the 4-momentum (which, as you recall, is naturally a covector). Thus the RHS can be considered as "the gravitational force", which does not look too different from the electromagnetic force $dp_\mu/d\tau = q(\partial_\mu A_\nu - \partial_\nu A_\mu)u^\nu$. However, there are two important differences. First, the mass m cancels: gravity is not a force but an acceleration. Note that, as expected, this acceleration is contained in the metric's first derivative $\partial_\mu g_{\nu\rho}$. The second difference is that the LHS and RHS of (6) do not by themselves make coordinate-independent sense! In particular, we've seen that $\partial_\mu g_{\nu\rho}$ can be made to vanish by an appropriate choice of coordinates. Nevertheless, we expect that the full equation (6) does make coordinate-invariant sense: the minimal action principle selects the longest possible worldline between two points, which in flat spacetime would be a straight line! The analogous notion in curved spacetime is called a geodesic, and eq. (6) is known as the geodesic equation. We will soon learn how to write it down properly.

Exercise 1. *As an intermediate step towards rewriting eq.* (6)*, let us put it in a form where* u^μ *has an upper index throughout. Prove that* (6) *is equivalent to:*

$$\frac{du^\mu}{d\tau} + g^{\mu\lambda}\left(\partial_\nu g_{\rho\lambda} - \frac{1}{2}\partial_\lambda g_{\nu\rho}\right)u^\nu u^\rho = 0 . \tag{7}$$

II. CONSERVATION LAWS; REDSHIFT; FRW METRIC

Consider again the equation of motion (6) of the free-falling particle. We can use it as a very clean example of Noether's Theorem: symmetries are associated with conservation laws. Suppose that the curved metric $g_{\mu\nu}(x)$ has a symmetry: there is some coordinate, e.g. x^1, on which the metric does not depend. Then we see immediately from eq. (6) that the momentum component $p_1 = mu_1$ conjugate to x^1 will be conserved throughout the motion! For the flat metric $\eta_{\mu\nu}$, this is of course true for all 4 coordinates (t, x, y, z) and their canonical conjugates – the components (p_t, p_x, p_y, p_z) of 4-momentum.

Exercise 2. *Consider a particle in flat spacetime, written in cylindrical coordinates* (t, ρ, ϕ, z)*, which are related to* (t, x, y, z) *via* $(x, y) = (\rho\cos\phi, \rho\sin\phi)$*. Which of the 4 coordinates* (t, ρ, ϕ, z) *does the metric depend on? Express the conserved quantity* mu_ϕ *in terms of the particle's position* $\mathbf{r} = (x, y, z)$ *and velocity* $\mathbf{v} = d\mathbf{r}/dt$*. What does this quantity represent?*

Of particular interest is the case of a time-independent metric $\partial_t g_{\mu\nu} = 0$. The technical term for such a metric is stationary. If we also have $g_{ti} = 0$, then the metric is called static; intuitively, in a static metric, the coordinate axes are "at rest", while in a stationary one, they can have a time-independent "velocity". For our present purpose, the stationarity condition $\partial_t g_{\mu\nu} = 0$ will be enough. For a free particle moving in the spacetime, it implies a conserved energy $-mu_t$ (note that this is no longer the same as mu^t!). Interestingly, this conserved energy isn't necessarily the same as what an observer sitting next to the particle may want to call energy. Indeed, in general, the "time axis vector" $\partial/\partial t = (1, 0, 0, 0)$ may not be normalized, i.e. g_{tt} may not be -1. For the local observer, it would make more sense to measure time not with t, but with a normalized time τ, which flows along the same direction as t, i.e.:

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{-g_{tt}}} \ ; \quad \frac{dx^i}{d\tau} = 0 \ . \tag{8}$$

The "energy" conjugate to such a time coordinate reads:

$$-mu_\tau = -m\frac{dt}{d\tau}u_t = -\frac{mu_t}{\sqrt{-g_{tt}}} \ , \tag{9}$$

which is not conserved, since g_{tt} can depend on x^i, and thus isn't constant along the particle's motion!

Exercise 3. *Show that, in the non-relativistic limit, the difference between* mu_t *and* mu_τ *amounts to taking or not taking into account the gravitational potential energy.*

There is another example of essentially the same effect, which on the surface might seem less esoteric. Energy is the derivative $-\partial S/\partial t$ of the action. We can similarly talk about the time frequency ω of e.g. a light wave, which is the time derivative $-\partial\phi/\partial t$ of the wave's phase (note that these two cases are essentially the same through quantum mechanics!). The propagation of a light wave through a time-independent metric is governed by a time-independent differential equation, which can be solved by separation of variables as $e^{-i\omega t}\psi(x^i)$. In other words, the frequency ω is conserved throughout the wave – another statement of Noether's theorem! However, now we can again consider the conserved $\omega = -d\phi/dt$ vs. the locally measured $\omega_{\text{proper}} = -d\phi/d\tau = \omega/\sqrt{-g_{tt}(x^i)}$. This dependence of the proper frequency, a.k.a. the color of the light, on the position x^i is known as gravitational redshift. We can understand it roughly in terms of kinetic vs. potential energy, as in Exercise 3: light climbing out of a potential well loses some of its "kinetic energy" $\hbar\omega_{\text{proper}}$, which manifests as a frequency shift towards the red.

Exercise 4. *On its way to us, sunlight is absorbed by atoms on the Sun's surface, at characteristic frequencies ω_{sun}. On Earth, we observe black lines in the Sun's spectrum at redshifted frequencies ω_{earth}. Find the redshift factor $1 - \omega_{earth}/\omega_{sun}$. As your inputs, you can use e.g. the Sun's radius, the Earth-Sun distance, and the length of a year.*

Another, complementary, example of symmetric metrics is when a metric is invariant along the spatial coordinates $\partial_i g_{\mu\nu} = 0$, but not along the time coordinate. Such a metric can be parameterized as:

$$ds^2 = -dt^2 + a(t)^2 \mathbf{dx}^2 \ , \tag{10}$$

where $a(t)$ is called the "scale factor" for obvious reasons: spatial distances are measured by $a(t)|\mathbf{dx}|$. The metric (10) is the known as the Friedman-Robertson-Walker (FRW) metric with flat spatial slices, and it is a good approximation for the large-scale structure of our (homogeneous, spatially flat, expanding) Universe. In this metric, there is no energy conservation. Instead, we have conserved spatial momentum $p_i = mu_i$ conjugate to the spatial coordinates x^i. For light waves, this corresponds to a conserved wavevector $\mathbf{k} = \partial\phi/\partial\mathbf{x}$. However, as before, this conserved wavevector isn't defined with respect to "proper" local coordinates which directly measure spatial distances. The locally measured proper wavevector is instead given by:

$$\mathbf{k}_{\text{proper}} = \frac{1}{a(t)}\frac{\partial\phi}{\partial\mathbf{x}} = \frac{\mathbf{k}}{a(t)} \ . \tag{11}$$

Thus, in an expanding Universe, the proper wavelength $\lambda_{\text{proper}} = 2\pi/|\mathbf{k}_{\text{proper}}|$ gets "passively stretched" as $\sim a(t)$, i.e. in the same way as the distances $a(t)|\mathbf{dx}|$. The light's proper momentum $\mathbf{p}_{\text{proper}} = \hbar\mathbf{k}_{\text{proper}}$ decays as $\sim 1/a(t)$, and so does its proper energy $E_{\text{proper}} = |\mathbf{p}_{\text{proper}}|$ (recall $E^2 - \mathbf{p}^2 = m^2 = 0$ for photons!). Thus, as the Universe expands, the energy of the radiation inside it decreases. This should be contrasted with the two other important components in the Universe's energy budget: non-relativistic matter, whose energy is mostly rest energy mc^2 and is conserved, and vacuum energy (a.k.a. the cosmological constant), which has a constant density, and thus increases as the Universe expands.

The cosmological metric (10) can also serve as a good demonstration of the effects of a curved metric on causality: the metric determines where the lightcones $ds^2 = 0$ lie, and those determine possible causal relationships between events. In the metric (10), it is easy to see that a lightray is given by:

$$dt^2 - a(t)^2 \mathbf{dx}^2 = 0 \quad \Longrightarrow \quad \left|\frac{\mathbf{dx}}{dt}\right| = \frac{1}{a(t)} \ . \tag{12}$$

Exercise 5. *Consider two versions of the FRW metric:*

- $a(t) \sim t^\alpha$ *for some positive power* α*. This metric has a Big Bang at* $t = 0$*, and is typical for a radiation-dominated or matter-dominated Universe.*

- $a(t) = e^{Ht}$ *with a positive parameter* H*, which is known as the Hubble constant. This metric extends through the entire range* $-\infty < t < \infty$*, and is typical of a vacuum-energy-dominated Universe.*

For each case, consider an observer at rest at $\mathbf{x} = 0$*. What is the earliest time (if any) and what is the latest time (if any) at which this observer can see a point at* $|\mathbf{x}| = r$*?*

III. DERIVATIVES IN CURVED SPACETIME ARE TRICKY; EXTERIOR DERIVATIVES

Let us go back to the free particle's equation of motion (6). We already convinced ourselves that its RHS makes little sense on its own. What about the LHS? There, we're trying to take a spacetime derivative (in this case, along the worldline) of a quantity with indices (in this case, a covector). It is a very important fact that, while the derivative $\partial_\mu f$ of a scalar makes a covector, the derivative ∂_μ of a tensor is not itself a tensor. Consider for example a vector v^μ, which transforms under coordinate redefinitions as:

$$v^\mu \ \rightarrow \ v'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu \ . \tag{13}$$

The derivative $\partial_\mu v^\nu$ then transforms as:

$$\begin{aligned}\frac{\partial v^\nu}{\partial x^\mu} \;\to\; \frac{\partial v'^\nu}{\partial x'^\mu} &= \frac{\partial}{\partial x'^\mu}\left(\frac{\partial x'^\nu}{\partial x^\rho} v^\rho\right) = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial}{\partial x^\sigma}\left(\frac{\partial x'^\nu}{\partial x^\rho} v^\rho\right) \\ &= \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\rho}\frac{\partial v^\rho}{\partial x^\sigma} + \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\sigma \partial x^\rho} v^\rho \ .\end{aligned} \tag{14}$$

The first term is the expected transformation rule for a tensor with one upper and one lower index. However, the second term spoils it. Similarly and slightly more simply, the derivative $\partial_\mu u_\nu$ of a covector transforms as:

$$\frac{\partial u_\nu}{\partial x^\mu} \;\to\; \frac{\partial u'_\nu}{\partial x'^\mu} = \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial u_\rho}{\partial x^\sigma} + \frac{\partial^2 x^\rho}{\partial x'^\mu \partial x'^\nu} u_\rho \ . \tag{15}$$

Exercise 6. *Prove this, starting from the covector transformation rule:*

$$u_\mu \;\to\; u'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} u_\nu \ . \tag{16}$$

The transformation rule for the derivative $\partial_\mu T^{\cdots}{}_{\cdots}$ of general tensors can be derived from these by constructing the tensor as a product of vectors and covectors. For each of the tensor's indices, we will get one unwanted term in the transformation law, as in (14) (for upper indices) or (15) (for lower indices).

Why is this happening? The problem is that when we take the derivative $\partial_\mu v^\nu$, we are comparing (or, in this case, subtracting) values of v^ν at two adjacent points, x^μ and $x^\mu + dx^\mu$. But the tangent spaces at these points are different! The coordinate basis, with respect to which we evaluate the components v^ν, is no longer the same when we move to an adjacent point!

There are some important exceptions to the fact that derivatives ∂_μ aren't tensorial. These exceptions all have to do with closed loops in spacetime, so that we're not really comparing components at different points. One such exception is the antisymmetrized derivative (or "curl") $\partial_{[\mu} u_{\nu]}$ of a covector. Indeed, we can see from (15) that the unwanted transformation term is symmetric in $\mu\nu$, and thus vanishes upon antisymmetrization. As we recall from our treatment of electromagnetism, what $2\partial_{[\mu} u_{\nu]}$ measures is the circulation $\oint u_\mu dx^\mu$ of u_μ along an infinitesimal closed loop. In fact, it's easy to check that any completely antisymmetrized derivative $\partial_{[\mu} T_{\nu\ldots\rho]}$ makes a tensor. These higher-order "curls" can similarly be interpreted in terms of higher-dimensional closed loops. The antisymmetrized derivative is so important that it has its own special name and notation – it is known as the exterior derivative d. It is no coincidence that this is the same d as in dx^μ. He who understands why will master the Universe.

IV. LIE DERIVATIVES AND THE "GROUP" OF COORDINATE TRANSFORMATIONS

Another important exception is the so-called Lie derivative. It can be understood as follows. Consider the naive, non-tensorial derivative $u^\nu \partial_\nu v^\mu$ of a vector v^μ along the vector u^μ. For the sake of this exercise, it is helpful to imagine both u^μ and v^μ as infinitesimal. The role of u^μ in the derivative $u^\nu \partial_\nu v^\mu$ is to drag the "tail" of the "arrow" v^μ from the point x^μ to an adjacent point $x^\mu + u^\mu$, and then compare this "dragged" $v^\mu(x^\nu)$ to the actual value at the new point, $v^\mu(x^\nu + u^\nu)$. The problem with this procedure, the reason why $u^\nu \partial_\nu v^\mu$ doesn't work as a tensor, is that nothing is telling us what to do with the "head" of v^μ's arrow! But what if u^μ was defined not only at x^μ, but also at the point $x^\mu + v^\mu$, where the "head" of v^μ is located? Then we'd know how to use u^μ to drag the both the tail and the head of v^μ! In particular, the tail will be dragged using $u^\mu(x)$, but the head will be dragged using $u^\mu(x^\nu + v^\nu) = u^\mu + v^\nu \partial_\nu u^\mu$. Overall the "dragged along u^μ" version of v^μ will read:

$$v^\mu_{\text{dragged}} = v^\mu + \delta(\text{head})^\mu - \delta(\text{tail})^\mu = v^\mu + (u^\mu + v^\nu \partial_\nu u^\nu) - u^\mu = v^\mu + v^\nu \partial_\nu u^\nu \ . \tag{17}$$

On the other hand, the value of v^μ at the new point $x^\mu + u^\mu$ is just $v^\mu + u^\nu \partial_\nu v^\mu$. Subtracting from this the dragged value, we obtain the so-called Lie derivative of v^μ along u^μ:

$$\mathcal{L}_u v^\mu = u^\nu \partial_\nu v^\mu - v^\nu \partial_\nu u^\mu \ . \tag{18}$$

Exercise 7. *Show that $\mathcal{L}_u v^\mu$ transforms correctly as a vector.*

We can also take Lie derivatives along u^μ of other tensor quantities. For a scalar, we define trivially $\mathcal{L}_u f = u^\mu \partial_\mu f$.

Exercise 8. *Demonstrate the Leibniz rule:*

$$\mathcal{L}_u (f v^\mu) = f \mathcal{L}_u v^\mu + v^\mu u^\nu \partial_\nu f \ . \tag{19}$$

Exercise 9. *By demanding the Leibniz rule for $\mathcal{L}_u(w_\mu v^\mu)$, derive the formula for the Lie derivative of a covector:*

$$\mathcal{L}_u w_\mu = u^\nu \partial_\nu w_\mu + w_\nu \partial_\mu u^\nu \ . \tag{20}$$

Exercise 10. *Use the Leibniz rule to derive a formula for the Lie derivative $\mathcal{L}_u T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n}$ of a general tensor.*

The geometric intuition for the Lie derivative of various tensors is essentially the same: we use a vector field $u^\mu(x)$, rather than just a vector u^μ at one point, to define a flow that drags both the "tails" and the "heads" of the relevant arrows. This can be expressed very simply if we choose a coordinate system adapted to u^μ, in which the components of u^μ are $(1,0,0,0)$ throughout spacetime. This is equivalent to choosing coordinates in which e.g. $u^\mu\partial_\mu = \partial/\partial x^0$, i.e. u^μ corresponds to one of the coordinate axes. In such adapted coordinates, all the $\partial_\mu u^\nu$ derivatives in (18)-(20) vanish, and the Lie derivative $\mathcal{L}_u$ just coincides with the naive derivative $u^\mu\partial_\mu$. To sum up, we can think of a Lie derivative as a flow along some coordinate. When taken together, such derivatives generate the symmetry of general coordinate transformations! In fact, we can recognize the Lie derivative of a vector (18) as the antisymmetric Lie bracket of this symmetry, i.e. the commutator between flowing along u^μ and flowing along v^μ:

$$\mathcal{L}_u v^\mu = -\mathcal{L}_v u^\mu \equiv [u,v]^\mu \ . \tag{21}$$

Exercise 11. *Show that flow along the vector field $[u,v]^\mu$ indeed behaves as a commutator between the flows along u^μ and v^μ:*

$$(\mathcal{L}_u\mathcal{L}_v - \mathcal{L}_v\mathcal{L}_u)w^\mu = \mathcal{L}_{[u,v]}w^\mu \ . \tag{22}$$

If u^μ and v^μ correspond to two axes $u^\mu = (1,0,0,0)$ and $v^\mu = (0,1,0,0)$ in the same coordinate system, then $\mathcal{L}_u$ and $\mathcal{L}_v$ must commute, since in these coordinates they become simply the partial derivatives ∂_0, ∂_1. Thus, the commutator $[u^\mu, v^\mu]$ measures the extent to which the vector fields $u^\mu(x)$ and $v^\mu(x)$ fail to describe axes of the same coordinate system!