General Relativity Fall 2018
Lecture 7: From special to general relativity

Yacine Ali-Haïmoud
(Dated: September 26, 2018)

From now we focus on *the* torsion-free, metric-compatible covariant derivative, with Christoffel connection.

PARTIAL-TO-COVARIANT SUBSTITUTION

The equivalence principle states that one can always chose a coordinate system in which the laws of physics are those of special relativity (SR) with no gravity. The covariant derivative is the mathematical tool that allows us to express this. Let us consider a few examples.

Stress-energy tensor conservation – In special relativity we saw that conservation of stress-energy implies $\partial_\mu T^{\mu\nu} = 0$. In GR, the same law holds in a locally-inertial coordinate system (LICS). In such a coordinate system, the Christoffel symbols vanish (they are linear in the first derivative of the metric), so $\partial_\mu = \nabla_\mu$. Hence, in a LICS, we have $\partial_\mu T^{\mu\nu} = \nabla_\mu T^{\mu\nu} = 0$. But this is a tensorial equation, so it holds in any coordinate system.

Partial-to-covariant rule – The same partial-to-covariant substitution applies e.g. to Maxwell's equations, which become $\nabla_\mu F^{\nu\mu} = 4\pi J^\nu$, $\nabla_{[\mu} F_{\nu\sigma]} = 0$. In a few cases, there may be subtleties, similar in spirit to the those arizing when quantizing a classical equation, due to the non-commutation of covariant derivatives (cf homework).

A tensorial equation like $\nabla_\mu T^{\mu\nu} = 0$ is said to be **generally covariant**, as opposed to just Lorentz-invariant: it holds in any coordinate system, not just for inertial, orthonormal coordinates.

GEODESICS

Geodesic equation – In special relativity the motion of a free-falling particle is such that $du^\mu/d\tau = 0$. We rewrite this as $u^\nu \partial_\nu u^\mu = 0$. Again, this is $u^\nu \nabla_\nu u^\mu = 0$ in a LICS, which is a tensor equality, hence must hold in any coordinate system. In other words, writing $u^\nu \nabla_\nu u^\mu = 0$ is equivalent to saying that there exist some coordinates (which we need not specify), in which free-falling particles move, instantaneously, on straight-line trajectories.

Geodesics as (local) extrema of proper time – Consider timelike paths (i.e. whose tangent vector is everywhere timelike), between two events A and B. We parametrize all the paths by $\lambda \in [0, 1]$, such that $P(\lambda = 0) = A$, $P(\lambda = 1) = B$. The proper time is

$$\tau = \int_0^1 d\lambda \sqrt{-g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} \equiv \int d\lambda \ \mathcal{L}(x^\mu, \dot{x}^\mu). \tag{1}$$

This is extremized for trajectories satisfying

$$\frac{d}{d\lambda}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu}\right) = \frac{\partial \mathcal{L}}{\partial x^\mu}, \tag{2}$$

$$\frac{d}{d\lambda}\left(\frac{g_{\mu\nu}\dot{x}^\nu}{\sqrt{-g_{\sigma\rho}\dot{x}^\sigma \dot{x}^\rho}}\right) = \frac{1}{2}\frac{1}{\sqrt{-g_{\sigma\rho}\dot{x}^\sigma \dot{x}^\rho}}(\partial_\mu g_{\gamma\delta})\dot{x}^\gamma \dot{x}^\delta. \tag{3}$$

We define by τ_* the proper time along the extremal trajectory. We multiply the equation by $d\lambda/d\tau_* = 1/\sqrt{-g_{\sigma\rho}\dot{x}^\sigma \dot{x}^\rho}$ and get

$$\frac{d}{d\tau_*}\left(g_{\mu\nu}\frac{dx^\mu}{d\tau_*}\right) = \frac{1}{2}(\partial_\mu g_{\gamma\delta})\frac{dx^\gamma}{d\tau_*}\frac{dx^\delta}{d\tau_*}. \tag{4}$$

Before continuing, let us notice that, if we define

$$\tilde{\mathcal{L}} \equiv \frac{1}{2} g_{\mu\nu}\frac{dx^\mu}{d\tau_*}\frac{dx^\nu}{d\tau_*}, \tag{5}$$

then we have found that

$$\frac{d}{d\tau_*}\left(\frac{\partial\tilde{\mathcal{L}}}{\partial(dx^\mu/d\tau_*)}\right)=\frac{\partial\tilde{\mathcal{L}}}{\partial x^\mu}, \tag{6}$$

which means that the trajectory extremizing proper time also extremizes

$$\int d\tau_*\tilde{\mathcal{L}}\equiv\frac{1}{2}\int d\tau_*\; g_{\mu\nu}\frac{dx^\mu}{d\tau_*}\frac{dx^\nu}{d\tau_*}. \tag{7}$$

So, although this is *not* the proper time, extremizing this is equivalent to extremizing the proper time, and it is easier to do so.

Moving on: expanding the $d/d\tau_*$, and multiplying by $g^{\nu\sigma}$, we find that the extremal solution satisfies the geodesic equation,

$$\frac{d^2x^\sigma}{d\tau_*^2}+\Gamma^\sigma_{\mu\nu}\frac{dx^\mu}{d\tau_*}\frac{dx^\nu}{d\tau_*}=0. \tag{8}$$

This is a useful trick to compute Christoffel symbols: start with the integral (7) and extremize it, instead of computing the Christoffel symbols directly.

Note that the geodesics obtained are only *local* maxima of proper time: there can be several geodesics between two points (think of the geodesics joining two non-antipodal points on a sphere!)

PARALLEL TRANSPORT

Definition – Consider a curve with tangent vector $T = d/d\lambda$. A vector X is said to be parallel-transported along the curve if $\nabla_T X = T^\mu\nabla_\mu X = 0$. To understand what this means, pick a LICS. This equation means that, locally, $T^\mu\partial_\mu X^\nu = 0$, i.e. the components of X *in the LICS* remain approximately constant, over small enough regions. This matches our intuitive idea of parallel-transporting a vector in flat spacetime: we simply keep its components in a cartesian coordinate system constant. This can be generalized if X is a tensor of arbitrary rank.

Geodesic equations are curves whose tangent vector is *parallel-transported along the curve*, $\nabla_U U = 0$.

Conservation of angles during parallel transport – Suppose the vector fields X and Y are parallel transported along a curve with tangent T. Then

$$\nabla_T(g_{\alpha\beta}X^\alpha Y^\beta)=\nabla_T(g_{\alpha\beta})X^\alpha Y^\beta+g_{\alpha\beta}\left(\nabla_T X^\alpha\;Y^\beta+X^\alpha\nabla_T Y^\beta\right)=\nabla_T(g_{\alpha\beta})X^\alpha Y^\beta=0, \tag{9}$$

by metric compatibility. So the norms and angles between vectors are conserved by parallel transport.

Parallel transport along finite paths – In flat spacetime, parallel-transporting a vector along different finite curves gives the same result. It is not so in curved spacetime! Consider the *example of a sphere.*

RIEMANN TENSOR

The Riemann tensor quantifies the *curvature* of spacetime. It also represents tidal fields.

Definition – Given any vector field V^α, $\nabla_{[\alpha}\nabla_{\beta]}V^\gamma$ is a tensor field. Let us compute its components in some coordinate system:

$$\begin{aligned}\nabla_{[\mu}\nabla_{\nu]}V^\sigma&=\partial_{[\mu}(\nabla_{\nu]}V^\sigma)-\Gamma^\lambda_{[\mu\nu]}\nabla_\lambda V^\sigma+\Gamma^\sigma_{\lambda[\mu}\nabla_{\nu]}V^\lambda\\&=\partial_{[\mu}(\partial_{\nu]}V^\sigma+\Gamma^\sigma_{\nu]\lambda}V^\lambda)+\Gamma^\sigma_{\lambda[\mu}\left(\partial_{\nu]}V^\lambda+\Gamma^\lambda_{\nu]\rho}V^\rho\right)\\&=\left(\partial_{[\mu}\Gamma^\sigma_{\nu]\lambda}+\Gamma^\sigma_{\rho[\mu}\Gamma^\rho_{\nu]\lambda}\right)V^\lambda\equiv\frac{1}{2}R^\sigma{}_{\lambda\mu\nu}V^\lambda.\end{aligned} \tag{10}$$

Since the left-hand side is a tensor field and V is a vector field, we conclude that $R^\sigma{}_{\lambda\mu\nu}$ is a tensor field as well. You can check that explicitly from the transformation law of Christoffel symbols.

This is the Riemann tensor, which measures the non-commutation of second derivatives of vector fields (remember that second derivatives of scalar fields do commute, by assumption). It is *completely determined by the metric*, and is *linear in its second derivatives*.

Expression in LICS – Let us compute the derivative of the Christoffel symbol in a LICS:

$$\partial_\mu \Gamma^\sigma_{\nu\lambda} = \frac{1}{2}\partial_\mu \left[g^{\sigma\delta}\left(\partial_\nu g_{\lambda\delta} + \partial_\lambda g_{\nu\delta} - \partial_\delta g_{\nu\lambda}\right)\right] = \frac{1}{2}\eta^{\sigma\delta}\left(\partial_\mu\partial_\nu g_{\lambda\delta} + \partial_\mu\partial_\lambda g_{\nu\delta} - \partial_\mu\partial_\delta g_{\nu\lambda}\right), \tag{11}$$

since the first derivatives of the metric vanish in a LICS. Therefore we get, in a LICS,

$$R_{\delta\lambda\mu\nu} = \frac{1}{2}\left(\partial_\lambda\partial_\mu g_{\nu\delta} - \partial_\lambda\partial_\nu g_{\mu\delta} - \partial_\delta\partial_\mu g_{\nu\lambda} + \partial_\delta\partial_\nu g_{\mu\lambda}\right). \tag{12}$$

Symmetries – The Riemann tensor $R^\sigma{}_{\lambda\mu\nu}$ is, by definition, (*i*) antisymmetric in its last two indices. From the expression in the LICS, we further see that the fully covariant tensor $R_{\sigma\lambda\mu\nu}$ is moreover (*ii*) antisymmetric in the first two indiecs, (*iii*) symmetric under exchange of the first and last pair, and (*iv*), satisfies the following identity:

$$R_{\sigma\lambda\mu\nu} + R_{\sigma\mu\nu\lambda} + R_{\sigma\nu\lambda\mu} = 0. \tag{13}$$

Although derived in a specific coordinate system, these symmetry properties are tensorial and remain true in any coordinate system.

Parallel transport along a finite closed loop – Let us consider a vector V defined at a point p of the manifold, and a small closed curve passing through p, with tangent vector $T = d/d\lambda$. We define the vector field W on the curve by parallel-transporting V, i.e. such that $W|_p = V$, and $\nabla_T W = 0$. We then ask what is W at p after being parallel-transported once around the curve.

By assumption, we have

$$0 = T^\nu \nabla_\nu W^\mu = T^\nu \partial_\nu W^\mu + T^\nu \Gamma^\mu{}_{\nu\sigma} W^\sigma = \frac{dW^\mu}{d\lambda} + \Gamma^\mu{}_{\nu\sigma}\frac{dx^\nu}{d\lambda}W^\sigma. \tag{14}$$

Let us pick an LICS centered at p (so p has coordinates 0 in this system). The Christoffel symbols vanish at p, but *only* there: elsewhere, they are small, but non-zero. To lowest order, $\Gamma^\mu{}_{\nu\sigma} \approx x^\gamma \partial_\gamma \Gamma^\mu{}_{\nu\sigma}|_p$, and $W^\sigma \approx V^\sigma$. We therefore have

$$W^\mu - V^\mu \approx -\partial_\gamma \Gamma^\mu{}_{\nu\sigma}|_p V^\sigma \int d\lambda \frac{dx^\nu}{d\lambda} x^\gamma = -\partial_\gamma \Gamma^\mu{}_{\nu\sigma}|_p V^\sigma \int d\lambda \frac{dx^\nu}{d\lambda} x^\gamma. \tag{15}$$

The last integral is antisymmetric: indeed, its symmetirc part is

$$\int d\lambda \left(\frac{dx^\nu}{d\lambda}x^\gamma + \frac{dx^\gamma}{d\lambda}x^\nu\right) = \int d\lambda \frac{d}{d\lambda}(x^\gamma x^\nu) = x^\gamma x^\nu|_p - x^\gamma x^\nu|_p = 0, \tag{16}$$

since we are integrating over a *closed* loop. Therefore, we find

$$W^\mu - V^\mu \approx \partial_{[\nu}\Gamma^\mu_{\gamma]\sigma} V^\sigma \int dx^{[\nu} x^{\gamma]} = \frac{1}{2} R^\mu{}_{\sigma\nu\gamma} V^\sigma \int dx^{[\nu} x^{\gamma]}, \tag{17}$$

where we used the expression for the Riemann tensor in a LICS (i.e. setting the Christoffel symbols to zero, but not their derivatives). The last integral is just the area enclosed by the curve. Take for example a circle in the x^1, x^2 coordinates, parameterized by $x^1 = r\cos\lambda$, $x^2 = r\sin\lambda$, with $\lambda \in [0, 2\pi]$. Then $x^{[1}dx^{2]} = \frac{1}{2}r^2 d\lambda$, and $\int x^{[1}dx^{2]} = \pi r^2$. You will work out examples in the homework. From a dimensional analysis, the Riemann tensor has dimensions of inverse length squared. The corresponding characteristic lengthscale can be seen as the radius of curvature of spacetime. So, qualitatively, we showed that $\Delta V \sim V \times (r/R)^2$, if r is the characteristic scale of the loop, and R the radius of curvature.