

General Relativity Fall 2019

Lecture 9: parallel transport; geodesics; general covariance

Yacine Ali-Haïmoud
September 30th 2019

PARALLEL TRANSPORT

Consider a curve $p(\tau)$ on a manifold \mathcal{M} , with tangent vector $\bar{V} \equiv d/d\tau$, with components $V^\mu = dx^\mu/d\tau$ in a coordinate basis. This is a vector field defined on the curve only. Now suppose we are given a vector $\bar{W}_{(0)}$ at some point $p(\tau=0)$ of the curve. We may define a vector field $\bar{W}(\tau)$ along the curve as follows: require that $\bar{W}(0) = \bar{W}_{(0)}$ and that $V^\alpha \nabla_\alpha W^\beta = 0$. Indeed, the latter condition implies

$$0 = V^\mu \nabla_\mu W^\nu = V^\mu \partial_\mu W^\nu + \Gamma_{\mu\sigma}^\nu V^\mu W^\sigma = \frac{dx^\mu}{d\tau} \frac{\partial W^\nu}{\partial x^\mu} + \Gamma_{\mu\sigma}^\nu V^\mu W^\sigma \Rightarrow \frac{dW^\nu}{d\tau} = -\Gamma_{\mu\sigma}^\nu V^\mu W^\sigma.$$

This is thus a first-order, linear ordinary differential equation, which, given initial conditions, has a unique solution. Note that there is **no guarantee that a single coordinate system covers all the curve**, but we can smoothly patch solutions in different coordinates. A vector W^α satisfying

$$V^\alpha \nabla_\alpha W^\beta = \nabla_{\bar{V}} W^\beta = 0$$

is said to be **parallel-transported along the curve with tangent vector V^α** .

Let us get an **intuitive understanding** of what this means. We said earlier that a manifold of dimension n can always be embedded into \mathbb{R}^d , with $d > n$, i.e. can be seen as a smooth curved hypersurface of \mathbb{R}^d . At a point $p(\lambda)$ on the curve, the vector $\bar{W}(\tau)$ belongs to the tangent space $\mathcal{V}_{p(\lambda)}$ of \mathcal{M} , which is a subspace of the tangent space of the embedding space \mathbb{R}^d . The latter is flat, so the notion of parallel-transporting a vector there is intuitive: pick a globally cartesian/inertial coordinate system, and move the vector while keeping its components constant. Parallel-transport \bar{W} in this way from $p(\tau)$ to $p(\tau + d\tau)$. This results in a new vector \tilde{W} at $p(\tau + d\tau)$. If \mathcal{M} is curved, in general **this vector does not belong to the tangent plane $\mathcal{V}_{p(\tau+d\tau)}$** . To obtain $\bar{W}(\tau + d\tau)$, **project this vector onto $\mathcal{V}_{p(\tau+d\tau)}$** – this is again well defined with a globally cartesian coordinate system. This series of operations can be shown to be **rigorously equivalent to what we defined as parallel-transport**.

Another way to understand this notion is by considering the **parallel-transport equation in a LICS**. The Christoffel symbols vanish in a LICS so the parallel transport equation is just $dW^\mu/d\tau = 0$, which corresponds to the usual notion of parallel-transporting a vector in flat space. The **Christoffel symbols only vanish at the origin of the LICS**, where the metric derivatives are exactly zero. Thus, one would have to redefine a LICS with every increment $d\tau$ if one wanted to use this intuitive parallel-transport definition. The covariant derivative does this automatically for you!

One can generalize the notion of parallel transport to tensors of arbitrary rank: a tensor T is said to parallel-transported along a curve with tangent vector V^α is $\nabla_{\bar{V}} T = V^\alpha \nabla_\alpha T = 0$.

We can now **understand the property $\nabla_\alpha g_{\beta\delta} = 0$ more intuitively**. Suppose two vectors X^α and Y^α are parallel transported along a curve with tangent vector V^α . The metric serves to compute their norms and angles between them. For instance $||\bar{X}||^2 = g_{\alpha\beta} X^\alpha X^\beta$ (remember that this can actually be positive, negative or null). This varies along the curve $p(\tau)$ according to

$$\frac{d}{d\tau} ||\bar{X}||^2 = V^\mu \partial_\mu ||\bar{X}||^2 = V^\mu \nabla_\mu ||\bar{X}||^2,$$

where we used the fact that $||\bar{X}||^2$ is a scalar, so its covariant derivative is equal to its partial derivative. Replacing by its explicit expression, and using Leibniz's rule, we get

$$\frac{d}{d\tau} ||\bar{X}||^2 = V^\mu \nabla_\mu (g_{\sigma\nu} X^\sigma X^\nu) = V^\mu (\nabla_\mu g_{\sigma\nu}) X^\sigma X^\nu + g_{\sigma\nu} (V^\mu \nabla_\mu X^\sigma) X^\nu + g_{\sigma\nu} (V^\mu \nabla_\mu X^\nu) X^\sigma = V^\mu (\nabla_\mu g_{\sigma\nu}) X^\sigma X^\nu = 0,$$

where we used the fact that X^α is parallel-transported along the curve. Thus we find that the **norms of vectors remain constant during parallel-transport**, and so do “angles” between vectors.

GEODESICS AGAIN

We already derived the geodesic equation (for timelike geodesics) in two ways. (i) By minimizing the proper time between two events, and (ii) by requiring that a free falling particle **instantaneously moves along a straight line at constant velocity in a LICS**. The geodesic equation is

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\nu\sigma}^\mu \frac{dx^\nu}{d\tau} \frac{dx^\sigma}{d\tau} = 0.$$

Now remember that the 4-velocity u^α is the tangent vector to a particle's worldline, with components $u^\mu = dx^\mu/d\tau$. So we may rewrite this equation as

$$\frac{du^\mu}{d\tau} + \Gamma_{\nu\sigma}^\mu u^\nu u^\sigma = 0 = u^\nu \nabla_\nu u^\mu.$$

This gives an elegant geometric definition: **a geodesic is a curve whose tangent vector is parallel-transported along itself**. This also allows to define the **acceleration 4-vector**:

$$a^\alpha \equiv u^\beta \nabla_\beta u^\alpha.$$

This vector vanishes along a geodesic, which is the trajectory of a free particle. However, in the presence of forces (for instance, electromagnetic forces on a charged particle), a particle will no longer move on a geodesic, but rather follow an equation of the form

$$m a^\alpha = m u^\beta \nabla_\beta u^\alpha = f^\alpha,$$

where m is the inertial mass and f^α is a **non-gravitational force vector field**.

Last but not least, we can similarly define **null geodesics**. These are curves $p(\lambda)$ with null tangent vector $\bar{P} = d/d\lambda$, which satisfy

$$P^\alpha \nabla_\alpha P^\beta = 0, \quad g_{\alpha\beta} P^\alpha P^\beta = P^\alpha P_\alpha = 0.$$

Note that we can always describe the exact same curve with another parameter λ' . Denote by $\bar{Q} \equiv d/d\lambda'$ the tangent vector in this new parametrization. It is related to \bar{P} as follows:

$$\frac{d}{d\lambda'} = \frac{d\lambda}{d\lambda'} \frac{d}{d\lambda} \Rightarrow Q^\alpha = \frac{d\lambda}{d\lambda'} P^\alpha \equiv f P^\alpha.$$

This immediately implies that Q^α is also null, as it should. Moreover,

$$Q^\alpha \nabla_\alpha Q^\beta = f P^\alpha \nabla_\alpha (f P^\beta) = f^2 P^\alpha \nabla_\alpha P^\beta + f \frac{df}{d\lambda} P^\beta = 0 + \frac{df}{d\lambda} Q^\beta.$$

Thus the most general equation of a null geodesic with tangent Q^α is that $\nabla_{\bar{Q}} \bar{Q} \propto \bar{Q}$. By appropriately parametrizing the curve, however, we can always set the proportionality coefficient to zero. In this case the parameter λ is called an **affine parameters**. Affine transformations $\lambda' = a\lambda + b$ still lead to the same equation $\nabla_{\bar{P}} \bar{P} = 0$.

GENERALLY COVARIANT EQUATIONS

Laws of physics are relations between **geometric objects**, i.e. objects defined independently of a basis, i.e. tensors. In special relativity, we restrict ourselves to inertial coordinate system, related to one another by Lorentz-transformations. We'll call objects which transform as tensors under Lorentz transformations **"Lorentz-covariant" or "Lorentz tensors"**, in contrast with **"generally covariant tensors"** which are bona fide tensors, and transform as such for general coordinate transformations.

We will use covariant derivatives to "uplift" special-relativistic equations (which are Lorentz-covariant) to generally covariant equations. The general recipe is as follows: start by writing an equation in a LICS. Write it in an explicitly "generally covariant tensorial form". This now becomes the general-relativistic equation. There are **some subtleties due to the non-commutation of covariant derivatives of vectors**, reminiscent of the subtleties in quantizing classical equations, due to the non-commutation of quantum operators. We'll see some examples once we introduce the Riemann tensors.

Let us see a few important examples.

Generally-covariant Dirac function

Under change of coordinates, the Dirac function changes as

$$\delta_D^{(4)}(x^\mu - a^\mu) = \left| \det \left[\frac{\partial x^{\nu'}}{\partial x^\nu} \right] \right| \delta_D^{(4)}(x'^{\mu'} - a^{\mu'}),$$

where by $[\partial x^{\nu'}/\partial x^\nu]$ I mean the **matrix** whose components are the partial derivatives.

Now, the metric components transform as

$$g_{\mu\nu} = \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^{\nu'}}{\partial x^\nu} g_{\mu'\nu'},$$

thus, seeing this as a matrix equation, we find

$$\det([g_{\mu\nu}]) = \left(\det \left[\frac{\partial x^{\sigma'}}{\partial x^\sigma} \right] \right)^2 \det([g_{\mu'\nu'}])$$

Since in a LICS, $g_{\mu\nu} = \eta_{\mu\nu}$ has determinant -1, it always has a negative determinant in any coordinate system.

We denote by $g \equiv \det([g_{\mu\nu}])$ the determinant of the metric components in a given coordinate system. From the two transformation rules above, we conclude that

$$\frac{1}{\sqrt{-g}} \delta_D^{(4)}(x^\mu - a^\mu) \quad \text{is invariant under general coordinate transformations.}$$

In other words, it behaves as a *scalar field*. In particular, since $\det(g) = -1$ in LICS, $\delta_D^{(4)}$ is **Lorentz-invariant**, i.e. invariant under Lorentz transformations (but it does need the $\sqrt{-g}$ factor to be generally covariant!).

4-current density

Consider particles (labeled by index n) with proper “charge” q_n – this could be electric charge, baryon number, etc... **For now we work in a LICS $\{x^\mu\}$.** In this LICS, suppose the particles have spatial coordinates $x_n^i(t)$. The density of charge is then

$$\rho(t, \vec{x}) = \sum_n q_n \delta_D^{(3)}(\vec{x} - \vec{x}_n(t)).$$

The charge current is

$$\vec{J}(t, \vec{x}) = \sum_n q_n \frac{d\vec{x}_n}{dt} \delta_D^{(3)}(\vec{x} - \vec{x}_n(t)).$$

We can group them together in a 4-current J^μ :

$$J^\mu(t, \vec{x}) = \sum_n q_n \frac{dx_n^\mu}{dt} \delta_D^{(3)}(\vec{x} - \vec{x}_n(t)).$$

This does not yet look Lorentz-invariant. We rewrite this as

$$J^\mu(x^\nu) = \sum_n q_n \int dt_n \frac{dx_n^\mu}{dt_n} \delta_D^{(4)}(x^\nu - x_n^\nu) = \sum_n q_n \int d\tau_n u_n^\mu \delta_D^{(4)}(x^\nu - x_n^\nu),$$

where u_n^μ is the 4-velocity of particle n .