General Relativity Fall 2019
Lecture 3: the geodesic equation

Yacine Ali-Haïmoud
September 10th 2019

NORMALIZATION OF 4-VELOCITY

Last lecture we defined the proper time τ of a massive particle, s.t. $d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu$, and the 4-velocity $\overline{u}$, with components $u^\mu = dx^\mu/d\tau$ in an arbitrary coordinate system. This implies the following normalization, valid in an arbitrary coordinate system:

$$\boxed{g_{\mu\nu} u^\mu u^\nu \equiv u^\mu u_\mu = -1}, \qquad [\textbf{notation:} \quad u_\mu \equiv g_{\mu\nu} u^\nu]$$

INVERSE OF COORDINATE TRANSFORMATION

Consider a general coordinate system $\{x^\mu\}$. It should be clear that

$$\frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu = \begin{cases} 1 & \text{if} \quad \mu = \nu \\ 0 & \text{otherwise} \end{cases} \quad .$$

Now, given another general coordinate system $\{x^{\mu'}\}$, we can use the chain rule to compute the partial derivative:

$$\frac{\partial x^\mu}{\partial x^\nu} = \frac{\partial x^\mu}{\partial x^{\alpha'}} \frac{\partial x^{\alpha'}}{\partial x^\nu} = \delta^\mu_\nu. \tag{1}$$

Thus the inverse of the matrix whose elements are $\partial x^\alpha / \partial x^{\alpha'}$ is the matrix whose elements are $\partial x^{\alpha'}/\partial x^\alpha$.

4-VECTORS

For now we **define a 4-vector as a geometric object $\overline{V}$ whose components transform like those of a 4-velocity**, i.e.

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu \qquad \text{[general coordinate change]}.$$

In particular, under a Lorentz transformation between two ICS,

$$V^{\mu'} = \Lambda^{\mu'}{}_\mu V^\mu \qquad \text{[Lorentz transformation between two ICS]}.$$

4-vectors (and their generalization, tensors), represent **observable, physical quantities**.

Given a 4-vector with components V^μ, we can define the object $\underline{V}$ with components

$$V_\mu \equiv g_{\mu\nu} V^\nu.$$

We will see later on that this object is a **dual vector**. Under a change of coordinates, we have

$$V_{\mu'} = g_{\mu'\nu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} \frac{\partial x^{\nu'}}{\partial x^\lambda} V^\lambda = \delta^\nu_\lambda \frac{\partial x^\mu}{\partial x^{\mu'}} g_{\mu\nu} V^\lambda = \frac{\partial x^\mu}{\partial x^{\mu'}} g_{\mu\lambda} V^\lambda = \frac{\partial x^\mu}{\partial x^{\mu'}} V_\mu,$$

where I used Eq. (1) to simplify the terms in red. In general, we define a **dual vector as a geometric object $\underline{V}$ whose components transform as above**. We will formalize these definitions in the upcoming lectures.

Although this is somewhat dated language, 4-vectors are sometimes referred to as **contravariant vectors** and dual vectors as **covariant vectors**.

MEASUREMENTS IN OBSERVER'S REST-FRAME

Now suppose we know the components $V^\mu = (V^0, \vec{V})$ of $\overline{V}$ in an ICS $\{x^\mu\}$, and want to compute them **in the rest-frame of some observer**, moving with 3-velocity $\vec{v}$ with respect to that ICS. Let us denote these components $V^\mu_{(\mathrm{obs})} = (V^0_{(\mathrm{obs})}, \vec{V}_{(\mathrm{obs})})$. They are related to V^μ by a **Lorentz boost**. In particular, the 0 component is given by

$$V^0_{(\mathrm{obs})} = \frac{1}{\sqrt{1-v^2}}\left(V^0 - \vec{v}\cdot\vec{V}\right).$$

Now recall that, in the initial ICS, the observer's 4-velocity has components $u^\mu = \frac{1}{\sqrt{1-v^2}}(1, \vec{v})$. Thus we have

$$V^0_{(\mathrm{obs})} = u^0 V^0 - \delta_{ij} u^i V^j = -\eta_{\mu\nu} u^\mu V^\nu = -g_{\mu\nu} u^\mu V^\nu = -u_\nu V^\nu,$$

where we used the fact that $g_{\mu\nu} = \eta_{\mu\nu}$ in an ICS. Now, the quantity $g_{\mu\nu} u^\mu V^\nu$ is **coordinate-independent**, as you can easily explicitly check for yourself. Thus we found a coordinate-independent way to write the zero-th component of $\overline{V}$ **measured by an observer**.

In the **homework**, you'll be asked to write a similar geometric expression for the spatial components of $\overline{V}$ in the observer's rest frame.

METRIC INVERSE

Given a metric $g_{\mu\nu}$ in a general coordinate system, we define the **inverse metric** with the same symbol, g, but with the indices up, $g^{\mu\nu}$. Specifically, we define it through the relation

$$g^{\alpha\mu} g_{\mu\beta} = \delta^\alpha_\beta = \begin{cases} 1 & \text{if} \quad \alpha = \beta \\ 0 & \text{if} \quad \alpha \neq \beta \end{cases}.$$

If we put the components of the metric and inverse metric in square matrices, the above equation just states that the product of the two matrices is the identity matrix. It should start being clear now that we only sum over repeated indices if one of them is up and the other is down (though we make exceptions for purely spatial indices, in a cartesian coordinate system). We will formalize the mathematical meaning of up and down indices in a couple of lectures.

Let's look at some examples. First, in an ICS, $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$. It should be clear the the inverse of this matrix has the same numerical components, so

$$\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1).$$

Let us now consider spherical polar coordinates $(x^0, x^1, x^2, x^3) = (t, r, \theta, \varphi)$, with line element

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2.$$

In these coordinates the metric has components

$$(g_{\mu\nu}) \equiv \begin{pmatrix} g_{00} & g_{01} & & \\ g_{10} & g_{11} & & \\ & & \ddots & \\ & & & \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2\theta \end{pmatrix}.$$

The components of the inverse metric are just the inverse of the matrix above, i.e.

$$(g^{\mu\nu}) \equiv \begin{pmatrix} g^{00} & g^{01} & & \\ g^{10} & g^{11} & & \\ & & \ddots & \\ & & & \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/r^2 & 0 \\ 0 & 0 & 0 & 1/(r^2\sin^2\theta) \end{pmatrix}.$$

Change of coordinates

Let us now write the definition of the inverse metric in primed coordinates:

$$g^{\alpha'\mu'} g_{\mu'\beta'} = \delta^{\alpha'}_{\beta'}.$$

We already know how the metric components change. Inserting this, we have

$$g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\beta}{\partial x^{\beta'}} g_{\mu\beta} = \delta^{\alpha'}_{\beta'}.$$

Now multiply both sides and simplify – make sure to check at every step that there are the **same free indices on each side**:

$$\begin{aligned}
g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{\partial x^{\beta'}}{\partial x^{\alpha}} g_{\mu\beta} &= \delta^{\alpha'}_{\beta'} \frac{\partial x^{\beta'}}{\partial x^{\alpha}} \\
g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \delta^{\beta}_{\alpha} g_{\mu\beta} &= \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \\
g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} g_{\mu\alpha} &= \frac{\partial x^{\alpha'}}{\partial x^{\alpha}}.
\end{aligned}$$

Let us now multiply by the inverse (unprimed) metric:

$$\begin{aligned}
g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} g_{\mu\alpha} g^{\alpha\beta} &= \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} g^{\alpha\beta} \\
g^{\alpha'\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \delta^{\beta}_{\mu} &= \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} g^{\alpha\beta} \\
g^{\alpha'\mu'} \frac{\partial x^\beta}{\partial x^{\mu'}} &= \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} g^{\alpha\beta}.
\end{aligned}$$

The last step is to multiply by $\partial x^{\beta'}/\partial x^{\beta}$. After simplification, this gives the transformation of the inverse metric under changes of coordinates:

$$\boxed{g^{\alpha'\beta'} = \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \frac{\partial x^{\beta'}}{\partial x^{\beta}} g^{\alpha\beta}}.$$

Again, you see that the priming-the-indices convention makes such transformation rules very easy to remember.

In the **homework**, you'll be asked to check this relation explicitly for spherical polar coordinates.

FREE TRAJECTORIES AS PATHS OF MAXIMUM PROPER TIME

We know that **in an ICS**, particles not subject to any forces move along **straight lines at constant velocity**, i.e. such that $v^i = dx^i/dt =$ constant. This mplies $v =$ constant, hence $dt/d\tau =$ constant, thus, $dx^\mu/d\tau =$ constant. Let us show that these trajectories correspond to paths of **maximum proper time**.

We denote events along the trajectory by $P(\lambda)$, where $\lambda \in [0, 1]$ is a parameter (so the set of events $\{P(\lambda); \lambda \in [0, 1]\}$ is the worldline). Consider trajectories starting at a given event $A = P(\lambda = 0)$ and ending at $B = P(\lambda = 1)$, with a timelike separation fron A. The total proper time along such a trajectory is, **in an ICS**,

$$\tau = \int \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu} = \int_0^1 d\lambda \sqrt{-\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}, \qquad \dot{x}^\mu \equiv \frac{dx^\mu}{d\lambda}.$$

Rewrite this as

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

Let us now search for the trajectories extremizing τ. They are solutions of the **Euler-Lagrange equations**:

$$\boxed{\frac{d}{d\lambda} \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = \frac{\partial \mathcal{L}}{\partial x^\mu}}. \tag{2}$$

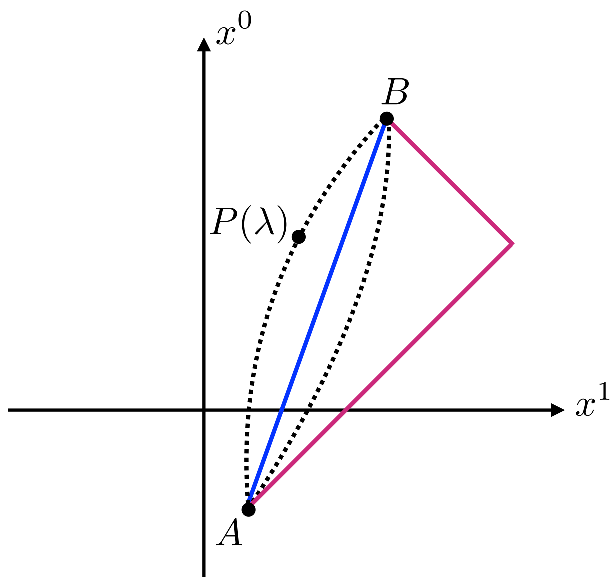

FIG. 1. Worlines between two events A and B, separated by a timelike separation. In an ICS, a particle not subject to any force moves on a straight line at constant velocity (in blue). We showed that this trajectory extremizes proper time. To show that it maximizes it, consider the purple paths, which is light-like (i.e. null) everywhere, hence as a vanishing proper time.

Computing each side, we get

$$\frac{d}{d\lambda}\left(-\frac{\eta_{\mu\nu}\dot{x}^\nu}{\sqrt{-\eta_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\right) = 0.$$

We recognize the denominator on the left-hand-side as $d\tau/d\lambda$. The Euler-Lagrange equations thus imply

$$\frac{d}{d\lambda}\left(\eta_{\mu\nu}\frac{dx^\nu}{d\tau}\right) = 0.$$

This implies that $\eta_{\mu\nu}\frac{dx^\nu}{d\tau}$ is constant, hence that $u^\mu = \frac{dx^\mu}{d\tau}$ is constant. Thus we have shown that the trajectories of extremal proper time are straight lines of constant velocity in ICS. It is easy to show that this is in fact the **maximum** proper time, as it is always possible to construct trajectories with null segments, hence with a zero proper time.

GEODESICS IN NON-INERTIAL COORDINATES

The conclusion we reached above is: **freely-moving particles move along paths of maximum proper time**. Even though we used knowledge about free-trajectories in ICS to arrive at this conclusion, it is clearly meaningful regardless of coordinates, given that proper time is coordinate independent. Let us now see what this translates to in a general coordinate system $\{x^\mu\}$. The proper time is given by

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

To compute the equation of motion in a general coordinate system, we look for extrema of τ, again using the Euler-Lagrange equations (2). This time, the right-hand-side is no longer zero, as $g_{\mu\nu}(x^\alpha)$ is a priori a function of the non-inertial coordinates. We thus get, for any $\mu \in \{0, 1, 2, 3\}$,

$$\frac{d}{d\lambda}\left(-\frac{g_{\mu\nu}\dot{x}^\nu}{\sqrt{-g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\right) = -\frac{1}{2}\frac{\partial_\mu g_{\rho\sigma}\ \dot{x}^\rho\dot{x}^\sigma}{\sqrt{-g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}, \qquad \left[\textbf{notation:}\ \ \partial_\mu X \equiv \frac{\partial X}{\partial x^\mu}\right].$$

Note that there isn't a 1/2 on teh right-hand-side because we need to take the derivative with respect to each of the two $\dot{x}^\mu$ factors. There are lots of indices here; note that all of them are repeated (hence summed over) except for μ. It is good practise to never have more than two indentical indices on each side to have absolutely no possible confusion on what is summed over. Let's now recall that $\sqrt{-g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta} = d\tau/d\lambda$, and rewrite the Euler-Lagrange equations as

$$\frac{d}{d\lambda}\left(\frac{g_{\mu\nu}\frac{dx^\nu}{d\lambda}}{\frac{d\tau}{d\lambda}}\right) = \frac{\partial_\mu g_{\rho\sigma}\ \frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda}}{\frac{d\tau}{d\lambda}}.$$

Multiply both sides by $d\lambda/d\tau = 1/(d\tau/d\lambda)$, and use the chain rule to simplify to

$$\frac{d}{d\tau}\left(g_{\mu\nu}\frac{dx^\nu}{d\tau}\right) = \frac{1}{2}\partial_\mu g_{\rho\sigma}\ \frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau}. \tag{3}$$

Let us now expand the left-hand-side:

$$\frac{d}{d\tau}\left(g_{\mu\nu}\frac{dx^\nu}{d\tau}\right) = \frac{d}{d\tau}(g_{\mu\nu}) \times \frac{dx^\nu}{d\tau} + g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2} = \partial_\alpha g_{\mu\nu}\frac{dx^\alpha}{d\tau}\frac{dx^\nu}{d\tau} + g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2},$$

where again we have used the chain rule to compute the derivative of the metric coefficients along the trajectory. We thus arrive at

$$g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2} = \frac{1}{2}\partial_\mu g_{\rho\sigma}\ \frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} - \partial_\alpha g_{\mu\nu}\frac{dx^\alpha}{d\tau}\frac{dx^\nu}{d\tau} = \left(\frac{1}{2}\partial_\mu g_{\rho\sigma} - \partial_\rho g_{\mu\sigma}\right)\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau},$$

where I used the fact that summation indices are dummy, in order to factorize the expression. Using the fact that the last factor is symmetric in ρ, σ, we can moreover replace $\partial_\rho g_{\mu\sigma} \to \frac{1}{2}(\partial_\rho g_{\mu\sigma} + \partial_\sigma g_{\mu\rho})$, and arrive at

$$g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2} + \frac{1}{2}\left(\partial_\rho g_{\mu\sigma} + \partial_\sigma g_{\mu\rho} - \partial_\mu g_{\rho\sigma}\right)\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0.$$

The last step is to multiply by $g^{\alpha\mu}$, the inverse metric. After simplifying, this gives

$$\boxed{\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0}, \tag{4}$$

where $\Gamma^\alpha_{\rho\sigma}$ is the **Chirstoffel symbol**:

$$\boxed{\Gamma^\alpha_{\rho\sigma} \equiv \frac{1}{2}g^{\alpha\mu}\left(\partial_\rho g_{\mu\sigma} + \partial_\sigma g_{\mu\rho} - \partial_\mu g_{\rho\sigma}\right) = \frac{1}{2}g^{\alpha\mu}\left(g_{\mu\sigma,\rho} + g_{\mu\rho,\sigma} - g_{\rho\sigma,\mu}\right)} \qquad [\textbf{notation: } X_{...,\nu} \equiv \partial_\nu X_{...}]$$

There aren't too many formulae that you should know "by heart" in physics, but this is one of them. Note that it is symmetric in the two lower indices. We have not yet defined formally what a tensor is, but one thing you should remember is that the Christoffel symbol is *not* a tensor.

Equation (4) is called the **geodesic equation**. This represents the equation of motion for a freely-moving particle in an arbitrary, non-inertial coordinate system. What it fundamentally means is that, in an ICS, a free-moving particle moves along a straight line with constant velocity, which also happens to be the path of maximum proper time. In a non-ICS, this path is described by a more complicated equation, even though it remains "fundamentally" (i.e. in an ICS) just a straight line traveled at constant velocity. In the homework, you'll explicitly derive the geodesic equation for spherical polar coordinates.

Other derivations and comments

• Let us first remark that we can arrive at the geodesic equation by extremizing the following integral (which does not have a particular physical meaning):

$$\mathcal{I} \equiv \int d\tau\ g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu, \qquad \dot{x}^\mu \equiv \frac{dx^\mu}{d\tau}.$$

Indeed, using the Euler-Lagrange equation with $\mathcal{L} = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$, we get precisely Eq. (3), from which the same steps follow. This is a **useful trick to derive the geodesic equation in an arbitrary coordinate system**, without first computing all the Christoffel symbols: just find the extrema of $\mathcal{I}$ using the Euler-Lagrange equation.

• Also note that **we can derive this equation starting from the geodesic equation in a ICS, and then changing coordinates**. In an ICS $\{x^\mu\}$, the geodesic equation is just

$$\frac{d^2x^\mu}{d\tau^2} = 0.$$

Now, using the chain rule, and the coordinate-independence of τ, we may rewrite this in terms of the general coordinates $x^{\mu'}$:

$$0 = \frac{d}{d\tau}\left(\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{dx^{\mu'}}{d\tau}\right) = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{d^2x^{\mu'}}{d\tau^2} + \frac{\partial^2 x^\mu}{\partial x^{\mu'}\partial x^{\nu'}}\frac{dx^{\mu'}}{d\tau}\frac{dx^{\nu'}}{d\tau}.$$

Recalling that the inverse of the matrix $\left(\partial x^\mu/\partial x^{\mu'}\right)$ is the matrix $\left(\partial x^\mu/\partial x^{\mu'}\right)$, we get

$$\frac{d^2x^{\alpha'}}{d\tau^2} + \frac{\partial x^{\alpha'}}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x^{\mu'}\partial x^{\nu'}}\frac{dx^{\mu'}}{d\tau}\frac{dx^{\nu'}}{d\tau} = 0.$$

This already has the form of the geodesic equation with

$$\Gamma^{\alpha'}_{\mu'\nu'} \equiv \frac{\partial x^{\alpha'}}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x^{\mu'}\partial x^{\nu'}}.$$

In the **homework**, you will show that this is indeed equal to the Christoffel symbol derived above.

• Once we introduce differential geometry and covariant derivatives, we will see that this equation can be written in a geometric way as

$$\frac{D}{D\tau}u^\alpha \equiv u^\beta \nabla_\beta u^\alpha = 0, \qquad u^\alpha \equiv \frac{dx^\alpha}{d\tau}.$$

The only difference between "full-on" GR and what we have been doing so far is that in general, ICS are only defined *locally*. The Christoffel symbol is proportional to derivatives of the metric components. Thus, the minimum requirements to *locally* recover the straight-line-at-constant-velocity free motion are not only $g_{\mu\nu} = \eta_{\mu\nu}$**, but also** $\partial_\alpha g_{\mu\nu} = 0$ – both of which are satisfied when we implicitly assume a *global* ICS, such that $g_{\mu\nu} = \eta_{\mu\nu}$ everywhere.

This gives us a definition of a **locally inertial coordinate system**: it is a coordinate system defined around a "central" event, around which the **metric components deviate from Minkowski only at second order in the coordinate separation**. In such a coordinate system, freely-moving particles travel *locally* (i.e. around the central event) on straight lines at constant velocity.