General Relativity Fall 2017
Lecture 2: The tangent space, coordinate bases

Yacine Ali-Haïmoud

September 7th 2017

This mostly mathematical lecture introduces concepts which wil be central to differential geometry and general relativity. The goal of this lecture is to formalize what is the underlying vector space which we use in Newtonian physics as well as special relativity, and to start thinking about vectors as living in a different space than the space(time) where material points (or events) live.

A. The Levi-Civita tensor [carry-over from Lecture 1]

The *Levi-Civita tensor* $\boldsymbol{\epsilon}$, which will be used to convey the notion of **volume**. This is a tensor of rank $(0, n)$ that is completely antisymmetric. The space of such tensors is of dimension 1: given a basis $e_{(1)}, ...e_{(n)}$ and the value $\epsilon_{12...n} \equiv \boldsymbol{\epsilon}(e_{(1)}, ..., e_{(n)})$, one can determine the value that $\boldsymbol{\epsilon}$ takes for any n-tuple. The only (possibly) non-zero components of ϵ are of the form

$$\epsilon_{\sigma(1)\sigma(2)...\sigma(n)} = s(\sigma)\epsilon_{12...n}, \tag{1}$$

where $\sigma \in \mathcal{P}_n$, and s is its signature.

The components of $\boldsymbol{\epsilon}$ change as follows for a change of basis:

$$\epsilon_{\mu'_1...\mu'_n} = \Lambda^{\mu_1}{}_{\mu'_1}\Lambda^{\mu_2}{}_{\mu'_2}...\Lambda^{\mu_n}{}_{\mu'_n}\epsilon_{\mu_1...\mu_n} = \sum_{\sigma\in\mathcal{P}_n} \Lambda^{\sigma(1)}{}_{\mu'_1}\Lambda^{\sigma(2)}{}_{\mu'_2}...\Lambda^{\sigma(n)}{}_{\mu'_n} s(\sigma)\epsilon_{12...n}. \tag{2}$$

So in particular,

$$\epsilon'_{1...n} = \sum_{\sigma\in\mathcal{P}_n} \Lambda^{\sigma(1)}{}_1\Lambda^{\sigma(2)}{}_2...\Lambda^{\sigma(n)}{}_n s(\sigma)\epsilon_{12...n}, \tag{3}$$

where I have changed the notation to put primes on tensors rather than indices, as it makes things clearer here. The bottom line is, the right-hand-side is just the definition of the determinant of Λ seen as a matrix:

$$\epsilon'_{12...n} = \det(\Lambda)\epsilon_{12...n}. \tag{4}$$

This tells us how one (hence all) components of $\boldsymbol{\epsilon}$ change under changes of coordinates.

The components of the metric change as follow:

$$g_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}g_{\mu\nu}. \tag{5}$$

Seen as a matrix equation, this is $g' = \Lambda^{\mathrm{T}} g\Lambda$. Taking the determinant of this relation, $\det(g') = \det(\Lambda^{\mathrm{T}} g\Lambda) = \det(\Lambda^{\mathrm{T}})\det(g)\det(\Lambda) = \det(\Lambda)^2\det(g)$. Therefore we find that the determinant of g must be of the same sign in any basis, and that

$$\det(\Lambda) = \pm\sqrt{\frac{\det(g')}{\det(g)}}. \tag{6}$$

So we conclude that

$$\frac{\epsilon'_{12...n}}{\sqrt{|\det(g')|}} = \mathrm{sign}\left[\det(\Lambda)\right]\frac{\epsilon_{12...n}}{\sqrt{|\det(g)|}}, \tag{7}$$

i.e., up to changes of sign, the ratio $\epsilon_{12...n}/\sqrt{|\det(g)|}$ is invariant under changes of basis.

We can now define THE Levi-Civita tensor, as the completely antisymmetric rank $(0, n)$ tensor whose component $\epsilon_{12...n} = 1$ in an orthonormal basis. The choice of that basis is not unique, an in particular we would get a sign flip for $\epsilon_{12\dots n}$ if we chose a basis with two flipped vectors. So the Levi-Civita tensor is only defined up to that arbitrary choice of *orientation*. Since in such a basis $\det(g) = \pm 1$ we then have, in any other basis,

$$\epsilon'_{12\dots n} = \mathrm{sign}\left[\det(\Lambda)\right] \sqrt{|\det(g')|}. \tag{8}$$

In summary, **the Levi-Civita tensor is the completely antisymmetric tensor of rank $(0, n)$ whose component $\epsilon_{12\dots n}$ is 1 in a "positively oriented" orthonormal basis**.

Exercise: show that in 3-dimensions, $\boldsymbol{\epsilon}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) \equiv \boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})$, which is the volume of the parallelepiped with sides $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$.

B. The spacetime interval, Lorentz transformations

Consider the usual flat physical space $\mathbb{R}^3$. If we pick cartesian coordinates (x, y, z), we know that the square distance between two points separated by $(\Delta x, \Delta y, \Delta z)$ is $\Delta \ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$. This distance is independent of the chosen cartesian coordinates (up to rescalings of course): if two coordinate systems (x, y, z) and (x', y', z') are related by a rotation, then $\Delta \ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 = \Delta x'^2 + \Delta y'^2 + \Delta z'^2$.

This notion gets generalized to the relativistic **spacetime interval** $\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$, where here and throughout we use units in which the speed of light $c = 1$. Let us review why this is the relevant interval. We recall the first fundamental postulate of special relativity: there exists coordinate systems (or reference frames) in which the speed of light $|d\vec{x}/dt| = c = 1$, independent of the velocity of the emitter. To this postulate we also add the assumption of **homogeneity** of spacetime (i.e., there are no preferred spacetime locations) and of **isotropy** of space (i.e., there are no preferred spatial directions).

Pick cartesian coordinates (x^1, x^2, x^3) [i.e. a system of rigid, orthogonal meterstics] on space, and a time $t \equiv x^0$, *defined* by synchronization through light signals, i.e. defined such that the time interval Δt it takes light to travel a distance $|\Delta \vec{x}|$ is such that $0 = -(\Delta x^0)^2 + (\Delta \vec{x})^2 \equiv \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$, where $\eta_{\mu\nu} \equiv \mathrm{diag}[-1, 1, 1, 1]$. Such a coordinate system is said to be **inertial** (we'll say more on this later on).

Now let us pick another inertial coordinate system (t', x', y', z'), i.e. by definition, a coordinate system for which $\eta_{\mu'\nu'} \Delta x^{\mu'} \Delta x^{\nu'} = 0$ along light paths, where again $\eta_{\mu'\nu'} \equiv \mathrm{diag}[-1, 1, 1, 1]$. By homogeneity of spacetime, the primed inertial coordinates must be linearly related to the unprimed inertial coordinates: $\Delta x^{\mu'} = \Lambda^{\mu'}{}_{\mu} \Delta x^\mu$ and vice-versa, $\Delta x^\mu = \Lambda^{\mu}{}_{\mu'} \Delta x^{\mu'}$. So we are looking for transformations $\Lambda^{\mu'}{}_{\mu}$ such that

$$0 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \Leftrightarrow 0 = \eta_{\mu'\nu'} \Delta x^{\mu'} \Delta x^{\nu'} = \eta_{\mu'\nu'} \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} \Delta x^\mu \Delta x^\nu. \tag{9}$$

where we emphasize that $\eta_{\mu'\nu'}$ has the same numerical values as $\eta_{\mu\nu}$. In matrix form, we are looking for transformatsions Λ such that

$$\Delta x^{\mathrm{T}} \eta \Delta x = 0 \Leftrightarrow \Delta x^{\mathrm{T}} (\Lambda^{\mathrm{T}} \eta \Lambda) \Delta x = 0. \tag{10}$$

This implies (though this is not completely trivial to prove) that $\Lambda^{\mathrm{T}} \eta \Lambda = \kappa \eta$, where κ is a constant, i.e.

$$-\Delta t'^2 + \Delta x'^2 + \Delta y'^2 + \Delta z'^2 = \kappa \left(-\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2\right). \tag{11}$$

Now we know that spatial rotations (with $t' = t$) satisfy this equation with $\kappa = 1$. Let us consider a boost, i.e. a change of coordinates such that the primed system has velocity $\boldsymbol{v}$ with respect to the unprimed system. By isotropy of space, the constant $\kappa(v)$ in Eq. (11) must only depend on the magnitude of $\boldsymbol{v}$. Now take three cordinate systems S_1, S_2, S_3, such that S_2 has velocity $\boldsymbol{v}_2$ wrt S_1, S_3 has velocity $\boldsymbol{v}_3$ wrt S_1, and S_3 has velocity $\boldsymbol{v}_{23}$ wrt S_2. Then we must have $\kappa(v_2)\kappa(v_{23}) = \kappa(v_3)$. Clearly, v_{23} must depend on the angle between $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ (it clearly does in the limit of small boosts). For the last equation to hold regardless of the angle, we therefore must have $\kappa = 1$, independently of the velocity. Hence we find that the transformations that preserve the condition $-dt^2 + d\vec{x}^2 = 0$ along light paths are such that

$$\Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} \eta_{\mu'\nu'} = \eta_{\mu\nu}. \tag{12}$$

This equation defines **Lorentz transformations**. A fortiori, this implies that the infinitesimal spacetime interval ds^2 is invariant under Lorentz transformations, i.e. independent of the inertial frame.

So the generalization of the Euclidean infinitesimal length squared $\Delta \ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$, which holds in cartesian coordinates and is invariant under rotations, is the Lorentzian spacetime interval $\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$, which holds for inertial coordinate systems, and is invariant under Lorentz transformations.

C. Tangent space, coordinate bases in flat spacetime

We are used to thinking of the flat spacetime as a vector space, since it is just $\mathbb{R}^4$. In fact however, we don't use the vector structure of spacetime itself: given two spacetime points P and Q, there is no meaning in $P+Q$ or $2P$. Spacetime (which we shall refer to as $\mathcal{M}$ for now on) is merely a set that "looks like" $\mathbb{R}^4$. We will formalize this when we talk about manifolds.

In fact the relevant vector space that we use for expressing physical laws is that of **infinitesimal displacements** near a spacetime event.

To convince yourself, think about the following: how does one calculate the length of a curve in $\mathbb{R}^3$, or equivalently the total spacetime interval along a curve in spacetime? We do so by adding the lengths of *infinitesimal* straight segments. Similarly, the velocity of a particle is an infinitesimal diplacement divided by the infinitesimal time it takes to travel across it.

We formalize the notion of "infinitesimal displacement along a direction X" with derivative operators. Define $\mathcal{F}$ as the set of differentiable functions $f : \mathcal{M} \to \mathbb{R}$. Given a spacetime point P and spacetime coordinates x^μ, we define the partial derivative operators

$$\partial_{(\mu)} : \begin{cases} \mathcal{F} \to \mathbb{R} \\ f \mapsto \partial_{(\mu)} f \equiv \frac{\partial f}{\partial x^\mu}\big|_P \end{cases} . \tag{13}$$

Suppose we are given a curve $Q(\lambda)$ on spacetime, where λ is some parameter, and $Q(0) = P$. The coordinates take the values $x^\mu(\lambda)$. Then the rate of change of a function f along this curve is obtained by the chain rule:

$$\frac{df}{d\lambda}\Big|_P = \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu}\Big|_P = \frac{dx^\mu}{d\lambda}\partial_{(\mu)} f. \tag{14}$$

This defines the operator $d/d\lambda \equiv \frac{dx^\mu}{d\lambda}\partial_{(\mu)}$. More generally, we define the **tangent space at** P, $\mathcal{V}_P$, as the space spanned by the $\partial_{(\mu)}$, i.e. for any vector $V \in \mathcal{V}_P$, $V = V^\mu \partial_{(\mu)}$. A vector V acting on a function simply gives its derivative along V. Even more generally, the tangent space is the space of linear operators on $\mathcal{F}$ satisfying Leibniz's rule, $V(fg) = f(P)V(g) + g(P)V(f)$, i.e. we could have defined it without introducing the operators $\partial_{(\mu)}$. The partial derivative operators $\partial_{(\mu)}$ form a basis of this space, called a **coordinate basis**: it is a special type of basis as it is attached to a coordinate system. We will soon give examples of non-coordinate bases of the tangent space.

From now on we will drop the parenthesis in the subscript: $\partial_\mu \equiv \partial_{(\mu)}$.

D. Dual vectors: gradients of functions

We can now use the formalism developed in Lecture 1, and define the dual vector space as the space of linear operators on vectors. In that case, the dual space is the space of **gradient of functions**: given a function $f : \mathcal{M} \to \mathbb{R}$, we define the operator df such that

$$df : \begin{cases} \mathcal{V}_P \to \mathbb{R} \\ V \mapsto V(f) \end{cases} . \tag{15}$$

This becomes a bit clearer in a coordinate basis: if $V = V^\mu \partial_\mu$, then $df(V) = V^\mu \partial f/\partial x^\mu$. In particular, if we take f as a coordinate x^ν, we get $dx^\nu(V) = V^\mu \partial x^\nu/\partial x^\mu = V^\mu \delta^\nu_\mu = V^\nu$. From this we see that the dx^ν are the dual basis of the coordinate basis ∂_μ: re-establishing the parenthesis for a minute, $dx^\nu(\partial_{(\mu)}) = \partial^\nu_{(\mu)} = \delta^\nu_\mu$. In the notation of Lecture 1, $dx^\mu \equiv \partial^{*(\mu)}$.

E. Flat spacetime metric

The metric tensor $\mathbf{g}$ conveys the notion of infinitesmal "squared distance" (which could be of either sign) per infinitesimal displacement along a vector V. Specializing to inertial coordinates in flat spacetime, we saw above that $\mathbf{g}(\partial_\mu, \partial_\nu) = \eta_{\mu\nu}$. This is equivalently expressed as $\mathbf{g} = \eta_{\mu\nu} dx^\mu \otimes dx^\nu$, often denoted by $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$, where ds^2 is the **line element**. We now know to think of this expression as the decomposition of the $(0,2)$ tensor on the dual basis of an inertial coordinate basis. In general (non-inertial) coordinates, the metric components need not be $\eta_{\mu\nu}$.

F. Change of coordinate basis

Suppose we have two coordinate systems $x^{\mu'}$ and x^μ and associated coordinate bases on the tangent space ∂_μ and $\partial_{\mu'}$. We can easily relate the two by using the standard chain rule:

$$\frac{\partial f}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial f}{\partial x^\mu}, \tag{16}$$

i.e.

$$\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu. \tag{17}$$

If we consider specifically inertial coordinate systems, which are related through one another through Lorentz transformations $\Delta x^\mu = \Lambda^\mu{}_{\mu'} \Delta x^{\mu'}$, in which case we simply have

$$\frac{\partial x^\mu}{\partial x^{\mu'}} = \Lambda^\mu{}_{\mu'}. \tag{18}$$

However, the rules for change of coordinates apply to any tensor acting on the tangent and dual space, i.e.

$$T^{\mu'_1 \dots \mu'_k}{}_{\nu'_1 \dots \nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x_{\mu_1}} \dots \frac{\partial x^{\mu'_k}}{\partial x_{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x_{\nu'_1}} \dots \frac{\partial x^{\nu_l}}{\partial x_{\nu'_l}} T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}. \tag{19}$$

As much as all this may seem like an overkill for now, it will be crucial when the spacetime is not flat!

Note that this notation makes intuitive sense: consider the dual basis vectors $dx^{\mu'}$ and dx^μ. Using the formalism of Lecture 1, they are related as

$$dx^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} dx^\mu, \tag{20}$$

i.e. exactly what one would expect if considering dx^μ as an infinitesimal change in x^μ rather than its formal definition of dual vector.

G. Example: spherical polar coordinates on $\mathbb{R}^3$

Let us look as spherical polar coordinates (r, θ, φ), defined such that

$$x = r \sin\theta \cos\varphi, \quad y = r \sin\theta \sin\varphi, \quad z = r \cos\theta. \tag{21}$$

The coordinates (r, θ, φ) describe physical space, and clearly don't have any vector space structure of their own.

Now, we just use the standard chain rule to express the partial derivative operators in terms of the cartesian partial derivative operators:

$$\partial_r = \frac{\partial x^\mu}{\partial r} \frac{\partial}{\partial x^\mu} = \sin\theta \cos\varphi \; \partial_x + \sin\theta \sin\varphi \; \partial_y + \cos\theta \; \partial_z, \tag{22}$$

$$\partial_\theta = \frac{\partial x^\mu}{\partial \theta} \frac{\partial}{\partial x^\mu} = r \cos\theta \cos\varphi \; \partial_x + r \cos\theta \sin\varphi \; \partial_y - r \sin\theta \; \partial_z, \tag{23}$$

$$\partial_\varphi = \frac{\partial x^\mu}{\partial \varphi} \frac{\partial}{\partial x^\mu} = -r \sin\theta \sin\varphi \; \partial_x + r \sin\theta \cos\varphi \; \partial_y. \tag{24}$$

We now compute the components of the metric on the dual basis $dr, d\theta, d\varphi$. All we need to do is to apply $\boldsymbol{g}$ to pairs of basis vectors. **Exercise:** Show the following results explicitly

$$g_{rr} = \boldsymbol{g}(\partial_r, \partial_r) = 1, \tag{25}$$

$$g_{r\theta} = 0, \tag{26}$$

$$g_{r\varphi} = 0, \tag{27}$$

$$g_{\theta\theta} = r^2, \tag{28}$$

$$g_{\theta\varphi} = 0, \tag{29}$$

$$g_{\varphi\varphi} = r^2 \sin^2\theta, \tag{30}$$

so the euclidean line element in these coordinates become

$$d\ell^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2, \tag{31}$$

which is the familiar spherical coordinate line element. **Exercise:** Show this result by using the coordinate transformation rule for tensors given the coordinate transformation of basis vectors.

This is a good point to emphasize that the metric is really a **tensor field**, in the sense that it is defined at each point of space (or spacetime). When expressed in cartesian coordinates, the metric coordinates were constant and this may not have been obvious, but it becomes clear in spherical coordinates, in which the metric components take on different values at different points.

This is also a good point to give an example of non-coordinate basis: instead of using $\partial_r, \partial_\theta, \partial_\varphi$, we could use the orthonormal basis

$$\hat{e}_r \equiv \partial_r, \quad \hat{e}_\theta \equiv \frac{1}{r}\partial_\theta, \quad \hat{e}_\varphi \equiv \frac{1}{r\sin\theta}\partial_\varphi, \tag{32}$$

defined everywhere except at the poles.

Exercise: show that there exists no coordinates for which this is a coordinate basis