General Relativity Fall 2019
Lecture 6: dual vectors and tensors

Yacine Ali-Haïmoud

September 19th 2019

COMMENTS ON HOMEWORK 2

• **Never ever have more than 2 repeated indices**. If you find that you have used more than two, go back and make sure you use different dummy indices.

• The primes on the indices fundamentally mean primes on the underlying object too. For instance, $g_{\mu'\nu'}$ means the components of the metric in the primed coordinate system. So $g_{\mu'\nu}$ is meaningless.

• **Partial derivatives associated with different coordinate systems do not commute:**

$$\frac{\partial}{\partial x^{\mu'}}\left(\frac{\partial f}{\partial x^{\mu}}\right) \neq \frac{\partial}{\partial x^{\mu}}\left(\frac{\partial f}{\partial x^{\mu'}}\right).$$

This holds in particular for $f = x^{\nu'}$: in general

$$\frac{\partial}{\partial x^{\mu'}}\left(\frac{\partial x^{\nu'}}{\partial x^{\mu}}\right) \neq \frac{\partial}{\partial x^{\mu}}\left(\frac{\partial x^{\nu'}}{\partial x^{\mu'}}\right) = \frac{\partial}{\partial x^{\mu}}(\delta^{\nu'}_{\mu'}) = 0.$$

Let us look for example at coordinates on $\mathbb{R}^2$, cartesian coordinates (x, y) on the one hand, and polar coordinates (r, ϕ) on the other hand, with $r = \sqrt{x^2 + y^2}$, $x = r\cos\phi, y = r\sin\phi$. Let's take for instance

$$f = \frac{1}{2}x^2 = \frac{1}{2}r^2\cos^2\phi.$$

We then have

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial r}\right) = \frac{\partial}{\partial x}(r\cos^2\phi) = \frac{\partial}{\partial x}\left(\frac{x^2}{r}\right) = \frac{x}{r}(1 + y^2/r^2),$$
$$\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial r}(x) = \frac{\partial}{\partial r}(r\cos\phi) = \cos\phi = \frac{x}{r}.$$

The reason the two do not commute is that the coordinates that are kept constant are different.

• **Instantaneous rest frame**: the frame in which, at a given instant, a particle is at rest, hence in this frame, its 4-velocity has components $(1, 0, 0, 0)$.

CHANGE OF BASES

Last lecture we defined the tangent space $\mathcal{V}_p$, and showed that it is a n-dimensional vector space, by proving that the partial derivative operators $\{\partial_{(\mu)}\}$ form a basis. This is called a **coordinate basis** of $\mathcal{V}_p$.

We will mostly use coordinate bases of the tangent space, but we don't have to do so. In general, we will denote bases of $\mathcal{V}_p$ with **subscripted and parenthesized indices**, e.g. $\{e_{(\mu)}\} = \{e_{(1)}, ..., e_{(n)}\}$, and the **components of a vector** $\overline{V}$ on a basis by V^μ, i.e. $\overline{V} = V^\mu e_{(\mu)}$. The parentheses are here to remind you that $e_{(\mu)}$ are *not* some components, but rather vectors. We place them down to make sure that **pairs of indices which are summed over always come with one up and one down**.

Now suppose that we have another basis $\{e_{(\mu')}\}$ (as before, putting the primes on the indices), related to the basis $\{e_{(\mu)}\}$ by

$$\boxed{e_{(\mu)} = M^{\mu'}{}_{\mu}e_{(\mu')}}.$$

In other words, **$M^{\mu'}{}_{\mu}$ is the μ'−th component of the vector $e_{(\mu)}$ on the basis $\{e_{(\mu')}\}$.** Then, given a vector $\overline{V}$, we can write

$$\overline{V} = V^\mu e_{(\mu)} = V^\mu M^{\mu'}{}_\mu e_{(\mu')} = V^{\mu'} e_{(\mu')}, \quad \boxed{V^{\mu'} = M^{\mu'}{}_\mu V^\mu}.$$

We see that to get the primed coordinates of $\overline{V}$ from its unprimed coordinates involes the primed-to-unprimed change-of-basis matrix. This is the origin of the word **"contravariant"** when talking about vectors.

Now let us apply this to **change of coordinate bases**. We first need to find the components of the vectors $\partial_{(\mu)}$ on the coordinate basis $\{\partial_{(\mu')}\}$. Last lecture, we saw that the components of a vector $\overline{V}$ on a coordinate basis are obtained by applying $\overline{V}$ to the coordinate functions x^μ, i.e. $V^\mu = \overline{V}(x^\mu)$. Thus, we have

$$M^{\mu'}{}_\mu = \partial_{(\mu)}(x^{\mu'}) = \frac{\partial x^{\mu'}}{\partial x^\mu}.$$

Thus we arrive at the following natural-looking relations:

$$\boxed{\partial_{(\mu)} = \frac{\partial x^{\mu'}}{\partial x^\mu}\partial_{(\mu')}, \qquad V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu, \quad \overline{V} = V^\mu \partial_{(\mu)}}.$$

EXAMPLE VECTOR: DERIVATIVE ALONG A CURVE

Consider a 1-dimensional smooth curve on $\mathcal{M}$ (the smoothness can be defined through charts) parametrized by some parameter $\tau \in \mathbb{R}$, i.e. the curve is a set of points $\{p(\tau)\}$. Given a function $f \in \mathcal{F}$, define the function

$$\tilde{f} : \begin{cases} \mathbb{R} & \to \mathbb{R} \\ \tau & \mapsto f(p(\tau)). \end{cases}$$

Then let us define the tangent vector $\overline{V} \equiv d/d\tau$ along the curve as follows: given a smooth function $f \in \mathcal{F}$, we define $d/d\tau(f) \equiv d\tilde{f}/d\tau$, where the right-hand-side is the usual derivative. It should be clear that $d/d\tau$ is a tangent vector. It should also be clear that its components on a coordinate basis $\{\partial_{(\mu)}\}$ are just $dx^\mu/d\tau$. We thus have

$$\frac{d}{d\tau} = \frac{dx^\mu}{d\tau}\partial_{(\mu)}.$$

The name τ was chosen on purpose to remind you of the proper time (which we will come back to once we formally introduce tensors). And indeed, the formal definition of the **4-velocity $\overline{u}$ is the tangent vector $d/d\tau$** along a curve parametrized by τ. It is a vector whose components on a coordinate basis are $dx^\mu/d\tau$.

DUAL VECTORS

A dual vector $\underline{W}$ at $p \in \mathcal{M}$ is defined as a **linear map $\underline{W} : \mathcal{V}_p \to \mathbb{R}$**. Note that this is a very general definition, that can be made for any vector space (i.e. it is not restricted to tangent spaces of manifolds). In words, a dual vector $\underline{V}$ acting on a vector $\overline{V}$ gives a real number, $\underline{W}(\overline{V}) \in \mathbb{R}$. Instead of parentheses, we will denote the action of dual vectors on vectors by a dot: $\underline{W}(\overline{V}) \equiv \underline{W} \cdot \overline{V}$. We define by $\mathcal{V}_p^*$ the set of dual vectors at some point $p \in \mathcal{M}$. I let you show for yourselves that this is a vector space.

Given a basis $\{e_{(\mu)}\}$ of $\mathcal{V}_p$, we can define the **dual basis $\{e^{*(\mu)}\}$ such that** $e^{*(\mu)} \cdot e_{(\nu)} = \delta^\mu_\nu$. I leave it as an **homework** exercise for you to show that this is indeed a basis of $\mathcal{V}_p^*$, which is thus also a vector space of dimension n. Given a vector $\overline{V} = V^\mu e_{(\mu)}$, we can easily show that its components are given by

$$V^\mu = e^{*(\mu)} \cdot \overline{V}.$$

Now consider again two bases of $\mathcal{V}_p$, related by $e_{(\mu)} = M^{\mu'}{}_\mu e_{(\mu')}$. For any vector $\overline{V}$, we have

$$e^{*(\mu')} \cdot \overline{V} = V^{\mu'} = M^{\mu'}{}_\mu V^\mu = M^{\mu'}{}_\mu e^{*(\mu)} \cdot \overline{V},$$

where we used the results from above. For this to hold for *any* vector $\overline{V}$, we must have

$$\boxed{e^{*(\mu')} = M^{\mu'}{}_{\mu} e^{*(\mu)}}.$$

Now, since $\{e^{*(\mu)}\}$ form a basis of $\mathcal{V}_p^*$, we can write any dual vector $\underline{W}$ as

$$\boxed{\underline{W} = W_\mu e^{*(\mu)}}.$$

The components are obtained by dotting the dual vector in the basis vectors:

$$W_\mu = \underline{W} \cdot e_{(\mu)}.$$

As you can suspect, and are asked to show explicitly, the components of a dual vector transform as

$$\boxed{W_\mu = M^{\mu'}{}_{\mu} W_{\mu'}},$$

that is, in the same way as the basis vectors. This is why dual vectors are also called **covariant vectors**.

Now, given a coordinate basis $\{\partial_{(\mu)}\}$, we can define its dual basis (just like for any basis!). We use the notation

$$\boxed{dx^{(\mu)} \equiv \partial^{*(\mu)}, \quad \text{i.e.} \quad \{dx^{(\mu)}\} = \text{dual basis of } \{\partial_{(\mu)}\}}.$$

From the transformation rule of the dual basis, we see that, under change of coordinates,

$$\boxed{dx^{(\mu')} = \frac{\partial x^{\mu'}}{\partial x^\mu} dx^{(\mu)}},$$

which looks exactly like the chain rule.

EXAMPLE DUAL VECTOR: GRADIENT OF A SCALAR FUNCTION

Consider a smooth function $f \in \mathcal{F}$. We define the dual vector $df \equiv \nabla f$ as follows:

$$\nabla f : \begin{cases} \mathcal{V}_p & \to \mathbb{R} \\ \overline{V} & \mapsto \overline{V}(f). \end{cases}$$

We denote the components of ∇f on a dual coordinate basis $\{dx^{(\mu)}\}$ by $\nabla_\mu f$:

$$\nabla f = \nabla_\mu f \ dx^{(\mu)}.$$

Now let us figure out the components on the coordinate basis, by computing $\nabla f \cdot \partial_{(\nu)}$:

$$\nabla_\mu f = \nabla f \cdot \partial_{(\nu)} = \partial_{(\nu)}(f) = \frac{\partial f}{\partial x^\nu}.$$

We have thus found

$$\boxed{df = \nabla f = \frac{\partial f}{\partial x^\mu} \ dx^{(\mu)}}.$$

The dual vector $df = \nabla f$ is called the **gradient of f**. Now consider a vector $\overline{V} = V^\mu \partial_{(\mu)}$:

$$\nabla f \cdot \overline{V} = V^\mu \nabla_\mu f = V^\mu \frac{\partial f}{\partial x^\mu}.$$

In other words, this gives the **directional derivative of f along $\overline{V}$**.

Note that the **expression above only holds in a coordinate (dual) basis**. The gradient ∇f is still well defined regardless of basis, but its components in a general basis are no longer $\partial f/\partial x^\mu$.

DUAL OF THE DUAL SPACE; CONTRACTION OF VECTORS AND DUAL VECTORS

Since $\mathcal{V}_p^*$ is a vector space, we can also define its dual $\mathcal{V}_p^{**}$, which is also a vector space of dimension n. Conveniently, there exists a **basis-independent bijective mapping between $\mathcal{V}_p^{**}$ and $\mathcal{V}_p$**: given a vector $\overline{V}$, define the following dual-dual vector (which by definition is a linear map from $\mathcal{V}_p^*$ to $\mathbb{R}$)

$$\tilde{V} : \begin{cases} \mathcal{V}_p^* & \to \mathbb{R} \\ \underline{W} & \mapsto \tilde{V} \cdot \underline{W} \equiv \underline{W} \cdot \overline{V}. \end{cases}.$$

I let you reflect on why the mapping $\overline{V} \mapsto \tilde{V}$ is bijective. The consequence is that we can see vectors as dual-dual vectors, i.e. **we can identify the dual-dual space with the original vector space, $\mathcal{V}_p^{**} \leftrightarrow \mathcal{V}_p$**. Note that there are many ways to build bijective mappings between $\mathcal{V}_p$ and $\mathcal{V}_p^*$ (an, in general, between any two vectors spaces of the same dimension), but there is no "geometric", basis-independent mapping as there is between $\mathcal{V}_p^{**}$ and $\mathcal{V}_p$.

We can thus define $\overline{V} \cdot \underline{W} \equiv \tilde{V} \cdot \underline{W} = \underline{W} \cdot \overline{V}$. This is called the **contraction** of the vector $\overline{V}$ and of the dual vector $\underline{W}$.

Now suppose that we have a basis $\{e_{(\mu)}\}$ of $\mathcal{V}_p$ and associated dual basis $\{e^{*(\mu)}\}$ of $\mathcal{V}_p^*$. Given a dual vector $\underline{W} = W_\mu e^{*(\mu)}$ and a vector $\overline{V} = V^\mu e_{(\mu)}$, we have

$$\underline{W} \cdot \overline{V} = W_\nu V^\mu e^{*(\nu)} \cdot e_{(\mu)} = W_\nu V^\mu \delta^\nu_\mu = W_\mu V^\mu.$$

Thus, **in any basis**,

$$\boxed{\underline{W} \cdot \overline{V} = \overline{V} \cdot \underline{W} = W_\mu V^\mu}.$$

In the expression $V^\mu W_\mu$, the repeated indices are callded **contracted indices**.

TENSORS

A rank (k, l) tensor at $p \in \mathcal{M}$ is a linear map

$$\boldsymbol{T} : (\mathcal{V}_p^*)^k \times (\mathcal{V}_p)^l \to \mathbb{R},$$

i.e. a linear operator that takes in k dual vectors $\underline{W}^{(1)}, ...\underline{W}^{(k)}$, and l vectors $\overline{V}_{(1)}, ..., \overline{V}_{(l)}$, and returns a real number

$$\boldsymbol{T}(\underline{W}^{(1)}, ...\underline{W}^{(k)}, \overline{V}_{(1)}, ..., \overline{V}_{(l)}) \in \mathbb{R}.$$

Note that the **parenthesized indices are labels**, and are **not components**. Just like we did for basis vectors and basis dual vectors, we use down labels for vectors and up labels for dual vectors, which works well with the **repeated-up-and-down-index summation convention**.

Given two dual vectors $\underline{X}, \underline{Y}$, we can construct the following tensor of rank (0, 2):

$$\underline{X} \otimes \underline{Y} : \begin{cases} \mathcal{V}_p \times \mathcal{V}_p & \to \mathbb{R} \\ (\overline{U}, \overline{V}) & \mapsto (\underline{X} \cdot \overline{U})(\underline{Y} \cdot \overline{V}). \end{cases}$$

This is called the **tensor product** of the dual vectors $\underline{X}, \underline{Y}$. Similarly, we can construct tensor products of arbitray numbers of vectors and dual vectors. For instance, given k vectors $(\overline{X}_{(1)}, ..., \overline{X}_{(k)})$ and l dual vectors $(\underline{Y}^{(1)}, ..., \underline{Y}^{(l)})$, we define

$$\overline{X}_{(1)} \otimes ... \otimes \overline{X}_{(k)} \otimes \underline{Y}^{(1)} \otimes ... \otimes \underline{Y}^{(l)} : \begin{cases} (\mathcal{V}_p^*)^k \times (\mathcal{V}_p)^l & \to \mathbb{R} \\ (\underline{W}^{(1)}, ...\underline{W}^{(k)}, \overline{V}_{(1)}, ..., \overline{V}_{(l)}) & \mapsto (\overline{X}_{(1)} \cdot \underline{W}^{(1)})...(\overline{X}_{(k)} \cdot \underline{W}^{(k)})(\underline{Y}^{(1)} \cdot \overline{V}_{(1)})... \times (\underline{Y}^{(l)} \cdot \overline{V}_{(l)}). \end{cases}$$

In particular, given a basis $\{e_{(\mu)}\}$ of $\mathcal{V}_p$ and a dual basis $\{e^{*(\nu)}\}$ of $\mathcal{V}_p^*$, we can construct a basis of the space of rank-(k, l) tensors with tensor products:

$$\{e_{(\mu_1)} \otimes ... \otimes e_{(\mu_k)} \otimes e^{*(\nu_1)} \otimes ... \otimes e^{*(\nu_k)}\},$$

where **each index** $\mu_1, ..., \mu_k, \nu_1, ..., \nu_l$ runs from 1 to n (or 0 to $n-1$). I let you think about why this is indeed a basis. This means that the space of tensors of rank (k, l) is of dimension n^{k+l}. We can then write a tensor as a linear combination of the basis vectors:

$$\mathbf{T} = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \ldots \nu_l} e_{(\mu_1)} \otimes ... \otimes e_{(\mu_k)} \otimes e^{*(\nu_1)} \otimes ... \otimes e^{*(\nu_l)},$$

where each repeated index is summed over (so this expression contains $k+l$ nested sums). The n^{k+l} real numbers $T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \ldots \nu_l}$ are the **components** of the tensor.

Note that we have already encountered tensors: **vectors are tensors of rank (1, 0)** and **dual vectors are tensors of rank (0, 1)**.

ABSTRACT INDEX NOTATION

So far we have been denoted vectors with a bar on top and dual vectors with a bar on the bottom. We could extend this notation to tensors but it would quickly become cumbersome. We could also adopt a slot notation, for instance, write $T^{\bullet\bullet\bullet}{}_{\bullet\bullet}$ to mean a rank (3, 2) tensor, but again, this is not super convenient.

So instead, we adopt the abstract index notation. We will say things like "consider the tensor $T^{\alpha\beta\gamma}{}_{\delta\lambda}$", meaning the **geometric** object $\mathbf{T}$ which is a rank (3, 2) tensor, **not** its components in a specific basis. In this context, the greek letters play the role of **placeholders**. I will try and keep greek letters close to the beginning of the alphabet to mean the geometric object, and further letters $\mu\nu, ...$ to mean the components, but the context should make meaning clear. All this will make more sense as we go and encounter specific examples.