General Relativity Fall 2019
Lecture 8: covariant derivatives

Yacine Ali-Haïmoud

September 26th 2019

METRIC IN NON-COORDINATE BASES

Last lecture we defined the metric tensor field $\boldsymbol{g}$ as a "special" tensor field, used to convey notions of infinitesimal spacetime "lengths". In a coordinate basis, we write $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ to mean $\boldsymbol{g} = g_{\mu\nu}\ dx^{(\mu)} \otimes dx^{(\nu)}$. While we will **mostly use coordinate bases**, we don't always have to. In a non-coordinate basis, we would write explicitly

$$\boldsymbol{g} = g_{\mu\nu}\ e^{*(\mu)} \otimes e^{*(\nu)}.$$

Let us consider for example flat 3-D space, in which the line element is

$$d\ell^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2$$

in cartesian and spherical polar coordinates, respectively. In homework 3, we assumed there existed some scalar product $\langle .. \rangle$ on the space of vectors; well, **this scalar product is just the metric: $\langle \overline{V}, \overline{W} \rangle \equiv \boldsymbol{g}(\overline{V}, \overline{W})$, and** $||\overline{V}||^2 \equiv \boldsymbol{g}(\overline{V}, \overline{V})$. We can read off the norms of the coordinate basis vectors from the line element:

$$||\partial_{(r)}||^2 = 1, \quad ||\partial_{(\theta)}||^2 = r^2, \quad ||\partial_{(\varphi)}||^2 = r^2 \sin^2\theta.$$

Thus the unit-norm vectors along the coordinate basis vectors are

$$e_{(r)} \equiv \partial_{(r)}, \quad e_{(\theta)} \equiv \frac{1}{r}\partial_{(\theta)}, \quad e_{(\varphi)} \equiv \frac{1}{r\sin\theta}\partial_{(\varphi)}.$$

The dual basis vectors are then

$$e^{*(r)} = dr, \quad e^{*(\theta)} = r\ d\theta, \quad e^{*(\varphi)} = r\sin\theta\ d\varphi.$$

Indeed, you can check explictly that these vectors satisfy $e^{*(\mu)} \cdot e_{(\nu)} = \delta^\mu_\nu$ – again, the "$\cdot$" operation represents the action of dual vectors on vectors, **not** the scalar product. In this **non-coordinate basis**, the line element is then

$$d\ell^2 \equiv \boldsymbol{g} = e^{*(r)} \otimes e^{*(r)} + e^{*(\theta)} \otimes e^{*(\theta)} + e^{*(\varphi)} \otimes e^{*(\varphi)}.$$

RAISING AND LOWERING INDICES WITH THE METRIC AND ITS INVERSE

We define the rank-(2, 0) **inverse metric tensor field** $g^{\alpha\beta}$ such that $g^{\alpha\beta} g_{\beta\gamma} = \delta^\alpha_\gamma$.

We saw last week that there isn't any basis-independent mapping from $\mathcal{V}_p$ to $\mathcal{V}_p^*$ (in contrast to the basis-independent mapping between $\mathcal{V}_p$ and $\mathcal{V}_p^{**}$). Now that we have a **special tensor** $g_{\alpha\beta}$, we can use it and its inverse to go from $\mathcal{V}_p$ to $\mathcal{V}_p^*$, without having to define a basis. Specifically, given a vector X^α, we define the dual vector X_α (using the same letter!)

$$\boxed{X_\alpha \equiv g_{\alpha\beta} X^\beta}.$$

Similarly, given a dual vector Y_α, we may define the vector Y^α

$$\boxed{Y^\alpha \equiv g^{\alpha\beta} Y_\beta}.$$

You can check that these two definitions are self-consistent. More generally, we can use the metric to define a rank $(k-1, l+1)$ tensor field from a rank-(k, l) tensor field, and the inverse metric to do the opposite. For instance, we have

$$T^{\alpha}{}_{\gamma}{}^{\beta}{}_{\delta} \equiv g_{\gamma\rho} g_{\delta\lambda} T^{\alpha\rho\beta\lambda}.$$

You see that the **position of the up and down indices matters**. For instance,

$$T^{\alpha\beta}{}_{\gamma\delta} - T^{\alpha}{}_{\gamma}{}^{\beta}{}_{\delta} = g_{\gamma\rho} g_{\delta\lambda} T^{\alpha\beta\rho\lambda} - g_{\gamma\rho} g_{\delta\lambda} T^{\alpha\rho\beta\lambda} = g_{\gamma\rho} g_{\delta\lambda} \left(T^{\alpha\beta\rho\lambda} - T^{\alpha\rho\beta\lambda} \right) \neq 0 \quad \text{in general,}$$

unless the rank-(4,0) tensor $T^{\alpha\beta\gamma\delta}$ is symmetric in its middle two indices.

COVARIANT DERIVATIVES

Given a **scalar field** f, i.e. a smooth function f – which is a **tensor of rank (0, 0)**, we have already defined the dual vector $\nabla_\alpha f$. We saw that, in a coordinate basis,

$$V^\alpha \nabla_\alpha f = V^\mu \frac{\partial f}{\partial x^\mu} \equiv \nabla_V f$$

gives the **directional derivative of f along V**.

We now want to generalize this idea of directional derivative to tensor fields of arbitrary rank, and we want to do so in a **geometric, basis-independent way**. We cannot just recklessly take derivatives of a tensor's components: **partial derivatives of components do not transform as tensors** under coordinate transformations. Indeed, given a vector field V^α, under a coordinate transformation, the partial derivatives of its components transform as

$$\frac{\partial V^{\mu'}}{\partial x^{\nu'}} = \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial}{\partial x^\nu}\left[\frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu\right] = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial V^\mu}{\partial x^\nu} + \frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\mu} V^\mu.$$

The presence of the second piece means that the partial derivatives of the components do not transform as a tensor. We thus need to correct **correct for this**: we have to find some **non-tensor coefficients $\Gamma^\mu_{\nu\sigma}$** such that

$$\nabla_\nu V^\mu \equiv V^\mu{}_{;\nu} \equiv \partial_\nu V^\mu + \Gamma^\mu_{\nu\sigma} V^\sigma$$

as a whole transforms as a tensor. These coefficients will thus transform as

$$\boxed{\Gamma^{\mu'}_{\nu'\sigma'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial x^\sigma}{\partial x^{\sigma'}}\Gamma^\mu_{\nu\sigma} + \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x^{\nu'}\partial x^{\sigma'}}}. \tag{1}$$

The proof that the combination $\partial_\nu V^\mu + \Gamma^\mu_{\nu\sigma} V^\sigma$ does indeed transform as a tensor is identical to Exercise 1 (*iii*) in homework 2.

• **Axiomatic definition**. We will define **a** covariant derivative in an axiomatic way – we will see that this definition does not uniquely specify **the** covariant derivative yet. Given a tensor $\boldsymbol{T}$ of rank (k, l), we define the tensor $\boldsymbol{\nabla T}$ of rank $(k, l+1)$, denoted as follows

$$(\nabla T)^{\alpha_1 \ldots \alpha_k}{}_{\beta_1 \ldots \beta_l \gamma} \equiv \nabla_\gamma T^{\alpha_1 \ldots \alpha_k}{}_{\beta_1 \ldots \beta_l} \equiv T^{\alpha_1 \ldots \alpha_k}{}_{\beta_1 \ldots \beta_l ; \gamma},$$

satisfying the following properties:

(*i*) the operator $\boldsymbol{\nabla}$ is **linear**, i.e. for two tensors $\boldsymbol{T}, \boldsymbol{S}$ of equal rank (and the same index structure), and two real numbers a, b, $\boldsymbol{\nabla}(a\boldsymbol{T} + b\boldsymbol{S}) = a\boldsymbol{\nabla T} + b\boldsymbol{\nabla S}$.

(*ii*) for a **scalar field** f, $\boldsymbol{\nabla} f$ is just the **usual gradient**.

(*iii*) it satisfies **Leibniz's rule**: given two tensors $\boldsymbol{T}, \boldsymbol{S}$ of arbitrary ranks (not necessarily equal),

$$\boldsymbol{\nabla}(\boldsymbol{T} \otimes \boldsymbol{S}) = \boldsymbol{\nabla T} \otimes \boldsymbol{S} + \boldsymbol{T} \otimes \boldsymbol{\nabla S},$$

(*iv*) it **commutes with contractions**: given a rank-(1, 1) tensor $T^\alpha{}_\beta$ - or more generally, a tensor of rank $(k \geq 1, l \geq 1)$,

$$\nabla_\delta (T^\alpha{}_\alpha) = \nabla_\delta (\delta^\alpha_\beta \; T^\beta{}_\alpha) = \delta^\alpha_\beta \; (\nabla T)^\alpha{}_{\beta\delta} = (\nabla T)^\alpha{}_{\alpha\delta}.$$

In words, we are requiring that the gradient of a scalar field (which is itself a contraction) is equal to the contraction of the rank-(1,2) tensor $(\nabla T)^\alpha{}_{\beta\gamma}$ in its first two indices, which is a non-trivial requirement.

• **Connection coefficients**. Let us now apply our axiomatic definition to the covariant derivative of a vector field. Suppose that we are given a coordinate basis $\{\partial_{(\mu)}\}$ that is **smoothly defined** around a neighborhood, so that each one of the basis vectors is a smooth vector field. The dual basis $\{dx^{(\mu)}\}$ will then also be defined around some neighborhood, and consist of smooth dual vector fields. Now consider a vector field $\overline{V} = V^\mu \partial_{(\mu)}$. The components of

$\overline{V}$ are $V^\mu = \overline{V} \cdot dx^{(\mu)}$, and are thus n **smooth scalar fields**. We can thus formally see the expression $\overline{V} = V^\mu \partial_{(\mu)}$ as a sum of tensor products of rank-(0,0) tensors (the components) with rank (1, 0) tensors: $\overline{V} = V^\mu \otimes \partial_{(\mu)}$.

Let us now compute a covariant derivative of $\overline{V}$, which is a rank (1, 1) tensor. Using Leibniz's rule, we get

$$\boldsymbol{\nabla}\overline{V} = \boldsymbol{\nabla}(V^\mu \otimes \partial_{(\mu)}) = (\boldsymbol{\nabla} V^\mu) \otimes \partial_{(\mu)} + V^\mu \otimes (\boldsymbol{\nabla}\partial_{(\mu)}).$$

By construction, any covariant derivative must just give the usual gradient when applied to scalar fields, thus $\boldsymbol{\nabla} V^\mu = \partial_\nu V^\mu dx^{(\nu)}$. Since $\boldsymbol{\nabla}\partial_{(\mu)}$ is a rank (1, 1) tensor, we may find its components on the basis $\{dx^{(\nu)} \otimes \partial_{(\sigma)}\}$:

$$\boxed{\boldsymbol{\nabla}\partial_{(\mu)} = \Gamma^\sigma_{\nu\mu}\ dx^{(\nu)} \otimes \partial_{(\sigma)}}.$$

This defines the **connection coeefficients $\Gamma^\sigma_{\nu\mu}$**. We thus found

$$\boldsymbol{\nabla}\overline{V} = \partial_\nu V^\mu dx^{(\nu)} \otimes \partial_{(\mu)} + \Gamma^\sigma_{\nu\mu} V^\mu\ dx^{(\nu)} \otimes \partial_{(\sigma)} = \left(\partial_\nu V^\mu + \Gamma^\mu_{\nu\sigma} V^\sigma\right) dx^{(\nu)} \otimes \partial_{(\mu)},$$

after renaming dummy indices. Thus we have

$$\boxed{\nabla_\nu V^\mu \equiv V^\mu{}_{;\nu} = \partial_\nu V^\mu + \Gamma^\mu_{\nu\sigma} V^\sigma}.$$

You see that the connection coefficients "connect" the covariant derivative to the partial derivative.

• **Covariant derivative of a dual vector field**. Consider a dual vector field W_α. For any vector field V^α, the contraction $V^\alpha W_\alpha$ is a scalar field. Thus, in a coordinate basis,

$$\nabla_\nu (V^\mu W_\mu) = \partial_\nu (V^\mu W_\mu) = (\partial_\nu V^\mu) W_\mu + V^\mu (\partial_\nu W_\mu),$$

per property (ii) of a covariant derivative, followed by Leibniz's rule for the usual partial derivative.

On the other hand, from property (iii) combined with (iv), we also have

$$\nabla_\nu (V^\mu W_\mu) = (\nabla_\nu V^\mu) W_\mu + V^\mu (\nabla_\nu W_\mu) = (\partial_\nu V^\mu) W_\mu + \Gamma^\mu_{\nu\sigma} V^\sigma W_\mu + V^\mu (\nabla_\nu W_\mu).$$

Equating the two, and renaming dummy indices, we get

$$V^\mu \left(\nabla_\nu W_\mu + \Gamma^\sigma_{\nu\mu} W_\sigma\right) = V^\mu \partial_\nu W_\mu.$$

This equality must hold **for any vector field V^α**. Thus we found the components of the rank-(0, 2) tensor $\boldsymbol{\nabla}\underline{W}$:

$$\boxed{\nabla_\nu W_\mu \equiv W_{\mu;\nu} = \partial_\nu W_\mu - \Gamma^\sigma_{\nu\mu} W_\sigma}.$$

Note the minus sign!

• **Covariant derivative of a general tensor field**. It is straightforward to generalize this to arbitrary tensor fields – you'll do this explicitly in the homework for rank-(0,2) and rank(1,1):

$$\boxed{\begin{aligned}\nabla_\sigma T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = \partial_\sigma T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} &+ \Gamma^{\mu_1}_{\sigma\lambda}\ T^{\lambda\mu_2\cdots\mu_k}{}_{\nu_1\cdots\nu_l} + \cdots \Gamma^{\mu_k}_{\sigma\lambda}\ T^{\mu_1\cdots\mu_{k-1}\lambda}{}_{\nu_1\cdots\nu_l}\\ &- \Gamma^\lambda_{\sigma\nu_1}\ T^{\mu_1\cdots\mu_k}{}_{\lambda\nu_2\cdots\nu_l} - \cdots \Gamma^\lambda_{\sigma\nu_l}\ T^{\mu_1\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_{l-1}\lambda}\end{aligned}}$$

In words, the covariant derivative is the partial derivative plus $k + l$ "corrections" proportional to a connection coefficient and the tensor itself, with a plus sign for all upper indices, and a minus sign for all lower indices.

THE TORSION-FREE, METRIC-COMPATIBLE COVARIANT DERIVATIVE

The properties that we have imposed on the covariant derivative so far are **not enough to fully determine it**. In fact, there is **an infinite number of covariant derivatives**: pick some coordinate basis, chose the $4^3 = 64$ connection coefficients in this basis as you wis. This thus defines the covariant derivative in this basis, hence defines in it any basis since $\boldsymbol{\nabla T}$ is a tensor – all you have to do is transform components from the original basis to any other basis. We will impose two more conditions on the covariant derivative to fully specify it.

• **Torsion-free**. From the transformation law of connection coefficients, you see that the non-tensorial part (the second derivative) drops out of the antisymmetric component:

$$\Gamma^{\mu'}_{[\nu'\sigma']} = \frac{\partial x^{\mu'}}{\partial x^{\mu}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial x^{\sigma}}{\partial x^{\sigma'}} \Gamma^{\mu}_{[\nu\sigma]}.$$

Thus, while $\Gamma^{\mu}_{\nu\sigma}$ **is not a tensor**, its antisymmetric part (in the lower two indices) $\Gamma^{\mu}_{[\nu\sigma]}$ **is indeed a tensor**. This is called the **torsion tensor field**. We will only consider covariant derivatives for which the torsion tensor field vanishes,

$$\boxed{\Gamma^{\mu}_{[\nu\sigma]} = 0}.$$

The most straightforward implication is the **commutation of double covariant derivatives of scalars**:

$$\nabla_\mu \nabla_\nu f - \nabla_\nu \nabla_\mu f = \partial_\mu(\nabla_\nu f) - \Gamma^{\sigma}_{\mu\nu} \nabla_\sigma f - \partial_\nu(\nabla_\mu f) + \Gamma^{\sigma}_{\nu\mu} \nabla_\sigma f = \partial_\mu(\nabla_\nu f) - \partial_\nu(\nabla_\mu f) + 2\Gamma^{\sigma}_{[\nu\mu]} \nabla_\sigma f.$$

Now $\nabla_\mu f = \partial_\mu f$ in a coordinate basis. Thus the two first terms cancel out, and we see that

$$\boxed{\nabla_\mu \nabla_\nu f - \nabla_\nu \nabla_\mu f = 2\Gamma^{\sigma}_{[\nu\mu]} \nabla_\sigma f = 0}.$$

• **Metric-compatible**. We want the covariant derivative to just recover the usual partial derivative in LICS. Thus we want to impose that $\Gamma^{\mu}_{\nu\sigma} = 0$ **in a LICS**. Note that if this is true in one LICS around $p \in \mathcal{M}$, it is true in all other LICS, which are just related by a Lorentz transformation. Indeed, for such transformations, the non-tensor part (the second derivatives) in the transformation (1) vanishes, since $x^{\mu'} = \text{const} + \Lambda^{\mu'}{}_{\mu} x^{\mu}$. This means that, if $\{x^{\mu'}\}$ is a LICS and $\{x^{\mu}\}$ is a general coordinate system, we have

$$\Gamma^{\mu}_{\nu\sigma} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial^2 x^{\mu'}}{\partial x^{\nu} \partial x^{\sigma}}, \qquad \{x^{\mu'}\} \text{ LICS}.$$

You already showed in homework 2 that this implies that the connection coefficients are just equal to the Christoffel symbols:

$$\boxed{\Gamma^{\mu}_{\nu\sigma} = \frac{1}{2} g^{\mu\lambda} \left(g_{\nu\lambda,\sigma} + g_{\sigma\lambda,\nu} - g_{\nu\sigma,\lambda} \right)}.$$

The fact that LICS are tied to the metric tensor ties the connection, hence covariant derivative to the metric tensor.

Another, equivalent way to arrive at the same conclusion, is to require that

$$\boxed{\nabla_\sigma g_{\mu\nu} = 0}.$$

You will show in the homework that this requirement indeed uniquely specifies the connection to be equal to the Christoffel symbols. We will see in a little what this requirement means in an intuitive way. The connection of the metric-compatible, torsion-free covariant derivative is also called the Levi-Civita connection.

Note: if we wanted to keep only metric-compatibility, but have a non-zero torsion, we could impose that $\Gamma^{\mu}_{(\nu\sigma)} = 0$ in a LICS. Equivalently, $\nabla_\mu g_{\nu\sigma} = 0$ only determines the symmetric part of the connection coefficients. So torsion-free and metric compatible are independent conditions on the antisymmetric and symmetric parts of the connection coefficients, respectively.