Lecture 4

In this lecture we talk about why the regular partial derivative does not form a tneosr, introduce the co-variant derivative and derive the Christoffel symbols.

12 The 4-Derivative

In any continuous theory we will of course expect the derivative to be an important object in the theory. In terms of the index notation we have been using so far we can write the traditional derivative using the following notation:

$$\partial_\mu := \frac{\partial}{\partial x^\mu}$$

We can see immediately see the reason for giving this quantity a lower index is that under a co-ordinate transformation $x^\mu \to y^\mu$, by the chain rule we have that:

$$\frac{\partial}{\partial x^\mu} \to \frac{\partial y^\nu}{\partial x^\mu}\frac{\partial}{\partial y^\nu} = \frac{\partial y^\nu}{\partial x^\mu}\partial'_\nu \Rightarrow \partial'_\nu = \frac{\partial x^\mu}{\partial y^\nu}\partial_\mu \tag{11}$$

That is, it transforms like a co-vector, hence the lower index. The presence of an index like this may give us the sense of security that this object, which by itself transforms like a tensor, is perfectly usable and will allow us to satisfy the principle of general co-variance. However, consider the following quantity. If we apply this 4-derivative to a 4-vector (contravariant) V^μ. Then this creates a scalar quantity, like the divergence of a vector field:

$$\partial_\mu V^\mu = \partial_0 V^0 + \partial_1 V^1 + \partial_2 V^2 + \partial_3 V^3$$

But if what happens if we transform co-ordinates from $x^\mu \to y^\mu$? Well V^μ transforms via 8 and ∂_μ via 11 resulting in:

$$\partial_\mu V^\nu \to \partial'_\mu (V')^\nu = \frac{\partial x^\alpha}{\partial y^\mu}\partial_\alpha\left(\frac{\partial y^\nu}{\partial x^\beta}V^\beta\right) = \frac{\partial x^\alpha}{\partial y^\mu}\frac{\partial y^\nu}{\partial x^\beta}\partial_\alpha V^\beta + \frac{\partial x^\alpha}{\partial y^\mu}\frac{\partial^2 y^\nu}{\partial x^\alpha x^\beta}V^\alpha$$

The first time on the RHS here looks exactly like the transformation that we would expect, but the second term completely rules out this quantity from being co-variant, despite the proper use of indices. This behavior of ∂_μ is not much of an issue in special relativity, as there the spacetime is flat and uniform, so there is not need for transformation of co-ordinates that are not linear. The most common transformation, a Lorentz transformation as seen in 3, is an example of such linear transformation, hence the second term in the above expression would vanish leaving a co-variant expression. This is equivalent to saying special relativity only deals with inertial frames.

13 The Co-variant Derivative

In GR however we have gathered that spacetime will need to curved and we want to be able to describe physics in non-inertia frames, hence we will ultimately need to need objects that are co-variant. This rules out ∂_μ as being a suitable candidate. The goal then will then be to create some object ∇ that is co-variant and reduces to the regular ∂_μ when space-time becomes flat. There are some properties of the derivative that this co-variant must satisfy, for us to consider it a derivative.

1. Linearity: $\nabla(\alpha V + \beta U) = \alpha\nabla V + \beta\nabla U$ (for constants α and β).
2. Leibniz rule: $\nabla(UV) = U(\nabla V) + (\nabla U)V$

These two properties make the derivative a linear operator. If we want the co-variant derivative to behave like a tensor we will also need to impose the following conditions:

3. Reduces to the partial derivative on Scalars: $\nabla_\mu \phi = \partial_\mu \phi$
4. Commutes with the contraction of indices: $\nabla_\mu\left(T^{\nu\gamma}{}_\gamma\right) = (\nabla T)_\mu{}^{\nu\gamma}{}_\gamma$

If we want to figure out the action of this operator on a vector we can first figure out how it acts on basis vectors. Recall from 5 that we denote our basis vectors b_μ. One thing we need to be careful of is how these vectors change at different points in space. When we move from one point to the next in a curved space time the space that the vectors in (known as the tangent space) can rotate and twist, making it very hard to compare vectors at different points. The process of deciding how to compare vectors at different points is not unique and we will have to make some choices to properly define it. For now we can simply assume that if at point p our basis vectors are given by b_μ then at the locally close point p' we can assume the basis at p', b'_μ is related to the basis at p by some matrix transformation:

$$b'_\mu = c^\nu{}_\mu b_\nu$$

If $p = p'$ then we would have that $b_\mu = b'_\mu \Rightarrow c^\nu{}_\mu = \delta^\nu_\mu$, so if p' is infinitesimally close to p, we should be able to expand the coefficients of the matrix $c^\nu{}_\mu$ as:

$$c^\nu{}_\mu = \delta^\nu_\mu + \epsilon \Gamma^\nu{}_\mu + \mathcal{O}(\epsilon^2)$$

Note that expanding to linear order will also work to satisfy the condition that the derivative be linear. With this we can go ahead and define the derivative in the traditional way:

$$\nabla b_\mu = \lim_{\epsilon \to 0} \frac{b'_\mu - b_\mu}{\epsilon} = \frac{(c^\nu{}_\mu - \delta^\nu_\mu) b_\nu}{\epsilon} = \Gamma^\nu{}_\mu b_\nu \tag{12}$$

But remember that the co-variant derivative has an index, as we can ask how the basis changes if we move in any direction. This means there is one of these matrices $(\Gamma_\rho)^\nu{}_\mu$ for each direction indexed by ρ. Often the following notation is used to refer to all these coefficients:

$$(\Gamma_\rho)^\nu{}_\mu := \Gamma^\nu_{\rho\mu}$$

These numbers are collectively known as the **Connection** as they connect vectors at different points in space, allowing them to be compared. With this definition, 12 becomes:

$$\nabla_\rho b_\mu = \Gamma^\nu_{\rho\mu} b_\nu \tag{13}$$

With this under out belt we can now figure out how the co-variant derivative acts on vectors $V = V^\mu b_\mu$:

$$\nabla_\rho V = \nabla_\rho (V^\mu b_\mu) = (\partial_\rho V^\mu) b_\mu + V^\mu \nabla_\rho b_\mu = (\partial_\rho V^\nu + V^\mu \Gamma^\nu_{\rho\mu}) b_\nu$$

Where we have used the fact that the components V^μ are just numbers, so using property 3: $(\nabla_\rho V^\mu) = \partial_\rho V^\mu$. What this equation tells us is that:

$$(\nabla_\rho V)^\nu = \partial_\rho V^\nu + V^\mu \Gamma^\nu_{\rho\mu} \tag{14}$$

That is, it tells us what the components of the co-variant derivative of any vector are. Now unfortunately, I have to break some bad news to you here. So far index notation had been our best of friend and has greatly simplified the notation we have been using. However, here it fails pretty bad. The common way to write the quantity we have just dealt with, namely $(\nabla_\rho V)^\nu$ is to forget the brackets and write $\nabla_\rho V^\nu$, notice this is **very** confusing, as I already told you that the co-variant derivative of the quantity V^μ alone is $(\nabla_\rho V^\mu) = \partial_\rho V^\mu$. But alas this is the notation that people have settled on. Just be wary that this is what most people mean. The same applies to the co-variant derivative of tensors.

Extending our above result 14 to finding the co-variant derivative of co-vectors and tensors is not too difficult if we consider the following argument. We know that the quantity $V^\mu U_\mu$ is a scalar meaning, by property 3 the co-variant derivative of this quantity must give us: $\nabla_\rho (V^\mu U_\mu) = \partial_\rho (V^\mu U_\mu)$, but we can also use Leibniz rule to expand:

$$\begin{aligned}\nabla_\rho(V^\mu U_\mu) = (\nabla_\rho V^\mu)U_\mu + V^\mu(\nabla_\rho U_\mu) = \partial_\rho(V^\mu U_\mu) = (\partial_\rho V^\mu)U_\mu + V^\mu(\partial_\rho U_\mu)\\ \Rightarrow (\partial_\rho V^\mu + V^\nu\Gamma^\mu_{\rho\nu})U_\mu + V^\mu(\nabla_\rho U_\mu) = (\partial_\rho V^\mu)U_\mu + V^\mu(\partial_\rho U_\mu)\\ \Rightarrow V^\mu(\nabla_\rho U_\mu) = V^\mu(\partial_\rho U_\mu) - V^\nu\Gamma^\mu_{\rho\nu}U_\mu\end{aligned}$$

So we have that the co-variant derivative acting on a co-vector has the following components:

$$\nabla_\rho U_\mu = \partial_\rho U_\mu - \Gamma^\mu_{\rho\nu}U_\mu$$

(Again what is really meant here is $\left(\nabla_\rho U\right)_\mu$). One can use this argument and extend it to find out how the co-variant derivative acts on all tensors. Doing this one finds:

$$\nabla_\rho T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_s} = \partial_\rho T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_s} + \sum_{a=1}^{k}\Gamma^{\mu_a}_{\rho\sigma}T^{\mu_1\ldots\sigma\ldots\mu_k}{}_{\nu_1\ldots\nu_s} - \sum_{a=1}^{s}\Gamma^{\sigma}_{\rho\nu_a}T^{\mu_1\ldots\sigma\ldots\mu_k}{}_{\nu_1\ldots\sigma\ldots\nu_s} \tag{15}$$

14 Christoffel Symbols

Now that we know the form of the co-variant derivative, we are left with some choices. The symbols[2] $\Gamma^\mu_{\rho\nu}$ are not unique given the 4 conditions we have imposed on the co-variant derivative so far. In order to obtain a unique connection we will impose the following two conditions:

5. The connection is torsion free: $\Gamma^\mu_{\rho\nu} = \Gamma^\mu_{\nu\rho}$
6. Metric compatibility: $\nabla_\rho g_{\mu\nu} = 0$

Note that this fifth condition is called torsion free, as one can define a Tensor known as the torsion tensor by the difference $T^\mu_{\rho\nu} = \Gamma^\mu_{\rho\nu} - \Gamma^\mu_{\nu\rho}$, and we are imposing that this tensor is 0. This is special to GR and there are actually other theories of gravity where this torsion tensor is non-zero and has a role to play.

We can then proceed to use condition 6 to find the form of the connection in GR. Using 15, and noticing that $\nabla_\rho g_{\mu\nu} = 0$ must be true for any permutation of the indices we can write:

$$\begin{aligned}\nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\rho\mu}g_{\sigma\nu} - \Gamma^\sigma_{\rho\nu}g_{\mu\sigma} = 0\\ \nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu}g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho}g_{\nu\sigma} = 0\\ \nabla_\nu g_{\rho\mu} = \partial_\nu g_{\rho\mu} - \Gamma^\sigma_{\nu\rho}g_{\sigma\mu} - \Gamma^\sigma_{\nu\mu}g_{\rho\sigma} = 0\end{aligned}$$

We can then add the first two equations together and take away the third:

$$\begin{aligned}\partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} - \Gamma^\sigma_{\rho\mu}g_{\sigma\nu} - \Gamma^\sigma_{\rho\mu}g_{\nu\sigma} - \Gamma^\sigma_{\mu\nu}g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho}g_{\nu\sigma} + \Gamma^\sigma_{\nu\rho}g_{\sigma\mu} + \Gamma^\sigma_{\nu\mu}g_{\rho\sigma} = 0\\ \Rightarrow 2\Gamma^\sigma_{\rho\mu}g_{\sigma\nu} = \partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu}\\ \Rightarrow \Gamma^\alpha_{\rho\mu} = \frac{1}{2}g^{\nu\alpha}(\partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu})\end{aligned} \tag{16}$$

Where we have used the torsion free property to cancel some of the $\Gamma^\sigma_{\nu\mu}$ symbols. Also in the last step we have applied the inverse metric to both sides in order to isolate $\Gamma^\alpha_{\rho\mu}$. This specific connection coeffecents we have derived are known as the **Christoffel symbols**, or the **Levi-Civita connection**. Note that as mentioned before they are not tensors, so one needs to calculate these symbols for every co-ordinate system seperatly.

As a sanity check we can try figure out what these symbols look like for flat spacetime, in this scenario, if we stick to cartesian co-ordinates the metric will be given by the Minkowski metric 4. Here the metric is independent of any of the co-ordinate variables, so all the derivatives in 16 vanish, meaning $\Gamma^\alpha_{\rho\mu} = 0$. This makes sense as the co-variant derivative reduces to the partial derivative as needed.

Note even if the space is flat there will still be some co-ordinates that have non zero Christoffel symbols, for example spherical co-ordinates in flat space time.

[2]Note that I am not calling this quantity a tensor as it can be shown that it does not transform like one, just like the partial derivative ∂_μ