[Lesson 14: The Covariant derivative of a covector]

Catalogue of Spacetimes: 1.3 Basic object of a metric
I want to prove these statements (1.3.3), I want to show how the connection on a manifold that models space-time is completely determined by the metric and under what conditions that's done. The typical way this is shown in General relativity books is through a calculus of variations method and I'm just not going to do that because it's so common, I'm trying to get there using an alternative way to show that some simple mathematical demands can also get us to this statement here there's the physical intuition that is shown in almost every General relativity textbook which is the basically the idea that your path through space-time should be extremal, it's a least action analysis, it's very very common and it's not very difficult although the math can sometimes get a little bit awkward if you're not really good with the comp notation of tensors which by now I think everybody here understands if you've made it this far but I want to be able to show that you can get this just with sheer mathematical reasoning on the model manifold itself and what the stipulations mathematically are that get us to this point. We're almost there, we've mastered some basic premises or basic expressions from the CFREE notation and now we're going to just get a few more and start pushing towards getting a derivation of this formula, it's worth the trouble.

The next thing we do have to show is, I want to show that the contraction operation commutes with the process of taking a derivative, the Covariant derivative ∇_X in a certain direction, this is a process and in the comp notation that would be $X^{\alpha}\nabla_{\alpha}Y^{\beta}$ and then you're taking the Covariant derivative of some other vector Y^{β} and you'll notice this ends up being a vector, this is a vector because normally this $\nabla_{\alpha}Y^{\beta}\equiv Y^{\beta}_{\|\alpha}$ or in more common textbook parlance that I don't like to use $\nabla_{\alpha}Y^{\beta}\equiv Y^{\beta}_{;\alpha}$. I've switched to this double bar notation. I want to show that contraction commutes with this operation of taking the Covariant derivative in a direction of a vector X .

To begin this we have to understand how contraction is written in CFREE notation, in regular notation and regular comp notation we would say the contraction of a tensor $X^{\alpha}{}_{\beta}$ is just $X^{\alpha}{}_{\alpha}$ but now we have to do something like the contraction operator on a tensor $C[X\otimes\alpha]$ you can only do the contraction if you have a tensor, a (p,q) rank tensor with neither p nor q are zero, this would be a $(1,1)$ tensor for example so the same $(1,1)$ tensor might look like this, it's just a tensor product of a vector and a co-vector which literally equals $\langle\alpha,X\rangle$. That's the CFREE idea of a contraction of X . It's a little bit ambiguous if I wanted to write a contraction of a more complex texture where I might have $C[X\otimes Y\otimes Z\otimes\alpha\otimes\beta\otimes\gamma]$ where α,β,γ are the co-vectors and X,Y,Z are vectors, then I have to specify which contraction I want, even the full contraction is a little bit ambiguous so you have to choose which guys are contracting with which, if I wanted to contract X with β I might have to write $C_{X\beta}$ operation there in which case I would get $Y\otimes Z\otimes\alpha\otimes\gamma\langle\beta,X\rangle$. For the first one it's pretty straightforward there's only one possible way to do it because there's only one vector and one co-vector and this is the basic rule for that circumstance:

$$C[X\otimes\alpha]=\langle\alpha,X\rangle \tag{1}$$

Now consider a tensor field $X(s)\otimes\alpha(s)$. This is a tensor field along some curve, the curve is γ the parameterization is τ at a certain value s so we get $\gamma(s)$ which is a point on the curve, in the manifold, it's the same thing as we always have but in this case I'm talking about this as a tensor field which means it's not auto parallel it's just a tensor field and this you could think of it as a function of space-time $X(x)\otimes\alpha(x)$ and that x is every point in the manifold, it can be, it's defined throughout the manifold or at least on in some neighborhood at some point on a manifold, it's defined continuously throughout this manifold. It's certainly if it's defined about the manifold it's certainly defined on this curve right and so $X(s)$ is now the value of this tensor field along this curve but it's not necessarily the parallel propagated value, it's just the value along the curve because we're talking about a tensor field now a real tensor field.

Now if I want to know the parallel propagated tensor relative to that field I would write $\Omega_s^{-1}[X\otimes S]_{\gamma(s)}$ so this would be how our notation works, this Ω_s Parallel transports a field from 0 to s but the inverse will take it from s to 0 so this thing is the Parallel transported version of this tensor field from s from $\gamma(s)$ the point in the manifold of $\gamma(s)$ to the point in a manifold of $\gamma(0)$ so that's what that symbol represents. Now we can ask the question, what about the contraction of that? We've already done the contraction here (1), what about the contraction of $\Omega_s^{-1}[X\otimes S]_{\gamma(s)}$? We use the facts that we generated in our last lesson where I will now write:

$$C\left[\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}\right]=C\left[\left(\Omega_s^{-1}X\right)\otimes\left(\Omega_s^{-1}\alpha\right)\right]=\left\langle\Omega_s^{-1}\alpha,\Omega_s^{-1}X\right\rangle \tag{2}$$

Because we are using the rule where the Parallel transport operator basically commutes with this tensor product operations, we get this expression but we know from this rule here (1) that the definition of contraction in the CFREE notation, the contraction of a tensor product of a vector and a co-vector is the inner product of the two. $\Omega_s^{-1}\alpha$ is the co-vector and $\Omega_s^{-1}X$ is the vector, the result is simply a number.

$$\left\langle\Omega_s^{-1}\alpha(s),\Omega_s^{-1}X(s)\right\rangle\Big|_0=\langle\alpha(s),X(s)\rangle \tag{3}$$

We know by our other rule that the dual space mapping of parallel propagated vectors and co-vectors is unchanged so we know that that is $\langle\alpha,X\rangle$ where this guy here is clearly evaluated at 0 because Ω_s would have been a mapping from 0 to s so Ω_s^{-1} is from s to 0 so these guys are here at s and these guys are here at 0 so those should be the same. If you Parallel transport them away from s it's the same as just taking that at s . That's an important thing about the contraction that we're gonna need so now I can apply the contraction to the difference of two specific tensors. I want to take the contraction and I want to apply it to this:

$$C\left\{\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}-[X\otimes\alpha]_{\gamma(0)}\right\} \tag{4}$$

Notice the first part is a tensor at $\gamma(s)$ and then this takes it from s to 0 so this whole thing is the tensor that is parallel to this field evaluated at s taken to 0 along a curve parameterize by s along the curve γ . The second part is that same tensor field evaluated at $\gamma(0)$. Notice these are not going to be the same, but this is the parallel propagated version of this tensor field at 0 from s parallel propagated along γ to 0 , that'll be some tensor at 0 but it won't be the same as that tensor field just evaluated at the point $\gamma(0)$, there is a difference here and of course they're two tensors are in the same point on the manifold therefore they're in the same tensor product space that'll be the tensor product space, $(1,1)$ tensor product space and at the manifold at $\gamma(0)$, it'll they'll both be in that tensor products base so you can subtract him you could also both multiply both of them by a number so the only number that's really relevant here is the parameter value s there's no other number in sight and if you divide both terms by the parameter s , you can do that because they're both tensors.

This is a tensor, this is a tensor, it lives in tensor product space, the tensor product space is a real vector space therefore you can multiply any of them by scalar so we're going to multiply it by $1/s$ and we're gonna get this object that's suspiciously beginning to look like a derivative right so I'm gonna ask myself well what about this contraction, this is very contrived but you'll see it's that's what proofs are proofs are guessing where to start and ending up at the right answer. Contractions are linear in vector sums and products, the contraction of the sum is the sum of the contractions that's a separate proof I think we might have done it in the past I don't know but it's easy to show because it's just adding up numbers you can go into the comp notation and show that pretty quickly. We're going to then conclude that this thing:

$$C\left\{\frac{\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}-[X\otimes\alpha]_{\gamma(0)}}{s}\right\}=\frac{1}{s}C\left\{\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}\right\}-\frac{1}{s}C\left\{[X\otimes\alpha]_{\gamma(0)}\right\} \tag{5}$$

We know how to evaluate these contractions, we're going to use this basic rule (1) that the contraction of a vector or the tensor is just the dual space mapping between the two so in the first case we get:

$$C\left\{\frac{\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}-[X\otimes\alpha]_{\gamma(0)}}{s}\right\}=\frac{1}{s}\langle\alpha(s),X(s)\rangle-\frac{1}{s}\langle\alpha(0),X(0)\rangle \tag{6}$$

It may look like you're doing things with tensors in different spaces but these are just numbers and we're just saying that this is the tensor product and the tensor sum, this is a tensor, that's a tensor that's a minus sign those are in the same tensor product space just so happens the numerical evaluation of this contraction I can use the value of the tensors at s and here I can use the value of the tensors at at 0 . I can use the value of the co-vector and vector at $\gamma(0)$. Now we just take this limit as $s\to0$. Now we consider that as a limit so if we consider that as a limit we look at this expression (left hand side of (6)) and consider as a limit in this expression (right hand side of (6)) and look consider it as a limit. Let's look at this one first, if we take s as a limit we end up with the contraction of:

$$C\left[\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\Omega_s^{-1}[X\otimes\alpha]_{\gamma(s)}\right]=\frac{1}{s}\left[\langle\alpha(s),X(s)\rangle-\langle\alpha(0),X(0)\rangle\right] \tag{7}$$

That's what this whole thing is as a limit. The definition of what this is, it's a derivative with respect to s , well that's got to be this thing as you shrink $s\to0$. I now have the contraction of this derivative and that I've just shown equals the right hand side of (6), what's this thing? Well that's a function and so this is $1/s$, this is the evaluation of a function at s and at 0 as $s\to0$, that's the just a straightforward derivative of that function and specifically because it's being evaluated as $s\to0$ along this curve γ , it really is going to be the definition of taking the derivative along this curve where every step of the way the vector along that curve the tangent vector is what we call $\dot{\gamma}$. The right side of this curve is of some function which is defined by this thing $\nabla_{\dot{\gamma}}$ in brackets here. Let me rewrite that as:

$$\frac{1}{s}\left[\langle\alpha(s),X(s)\rangle-\langle\alpha(0),X(0)\rangle\right]=\frac{\langle\alpha,X\rangle_{\gamma(s)}-\langle\alpha,X\rangle_{\gamma(0)}}{s} \tag{8}$$

That's a little bit easier to understand because now you see that this term here is being evaluated along the curve $\gamma(s)$ so as $s \to 0$ this thing crawls along the curve from $\gamma(s)$ to $\gamma(0)$ and it moves down this curve and you're evaluating this expression and each point along this curve. That makes it easy to see that this is the Covariant derivative in the direction of the tangent vector of this function. It's a function minus a number but that thing here that difference is always going to be the contraction of $X \otimes \alpha$ and so that shows us, now we just put these things together, this contraction here by definition is what this is right here in the middle $C(X \otimes \alpha)$ and this combined with the derivative and the fact that you're evaluating it in this curve and that this is going to 0 that's what creates this structure $\nabla_{\dot{\gamma}}$:

$$\nabla_{\dot{\gamma}}[C(X \otimes \alpha)] = C \nabla_{\dot{\gamma}}(X \otimes \alpha) \tag{9}$$

That shows that the contraction operator and the partial derivative in the direction of our curve commute. Important point that those two things commute. Why did we do that? What we're after is a comp expression for $\alpha_{\kappa \| \delta}$, the Covariant derivative of a 1-form α , we want that expression, we've already got an expression for the Covariant derivative of a vector $X^{\kappa}{}_{\| \delta}$. We know that that's:

$$X^{\kappa}{}_{\| \delta} = X^{\kappa}{}_{| \delta} + \Gamma^{\kappa}_{\alpha \delta} X^{\alpha} \tag{10}$$

I think that's it and so we want an equivalent expression for $\alpha_{\kappa \| \delta}$. Spoiler alert, it involves a minus sign here and obvious rearrangement of the indices over here but otherwise it's very similar, but we've got to prove it, we've got to get there so that's our goal so with that in mind, we can get there almost entirely using CFREE analysis based on the just the algebra we've learned so let's begin by taking the contraction of the Covariant derivative in the direction of X of a tensor:

$$C\{\nabla_X(Y \otimes \alpha)\} = C\{\nabla_X Y \otimes \alpha + Y \otimes \nabla_X \alpha\} \tag{11}$$

We used the Leibniz rule on the inside and then the linearity of the contraction. We haven't explored $\nabla_X \alpha$ that much, we don't know much about the Covariant derivative in the direction of X of a 1-form, we understand $\nabla_X Y$ real well, we can go into comp notation, if we had to go to comp notation right now we would have no trouble except for $\nabla_X \alpha$ but we know everything else, it's going to turn out that we understand how to do the comp notation for everything except $\nabla_X \alpha$. That's leaves us a hint of what to do we expand this as much as we can we put this on the left hand side of an equation then convert everything on the right hand side to comp notation and that will be the comp expression for this so that's our that's our strategy here. We're just going to keep expanding like crazy. On the left side I can swap the contraction with the Covariant derivative because we just proved that you could do that

$$\nabla_X[C(Y \otimes \alpha)] = \nabla_X \langle \alpha, Y \rangle = \langle \alpha, \nabla_X Y \rangle + \langle \nabla_X \alpha, Y \rangle \tag{12}$$

On the Right side, the contraction of the sum of two tensors is the sum of the contractions that's sort of an elementary fact about contractions. Remember $\nabla_X Y$ is still a vector right this is a $(1,1)$ tensor which in comp notation is $Y^{\alpha}{}_{\| \delta}$, that's a $(1,1)$ tensor but when you contract that against X you get $Y^{\alpha}{}_{\| \delta} X^{\delta}$, δ contracts so you're left with one index α so it's a vector. You have a vector and a 1-form or a vector and a co-vector.

We're counting on $\nabla_X \alpha$ to be a 1-form just like $\nabla_X Y$ this is a vector and the reason it has to be that way is the way we've defined contraction in CFREE notation and the way we've proven. We've proven that contraction commutes with with Covariant derivatives in a certain directions, we've proven that the Covariant derivatives in the direction of X is a Leibniz thing and we know that contraction is linear over the sums of tensors so it must be the you have to be able to execute this contraction and therefore that's the expression we're ultimately going to hunt down we can also simplify this side and we get:

$$\langle \nabla_X \alpha , Y \rangle = \nabla_X \langle \alpha , Y \rangle - \langle \alpha , \nabla_X Y \rangle \tag{13}$$

We're getting close, we're isolating we're slowly isolating this guy he's still trapped in this dual space mapping but that's not too much of a problem so what we'll do is we'll open up this dual space mapping and we will start expressing things in comp notation so now we're going to go into comp notation so let's start with this guy $\nabla_X \langle \alpha , Y \rangle$, this is a function so this is the Covariant derivative in the direction of X of a function and we know that the Covariant derivative in the direction X of a function is X operating on that function $\nabla_X \mathrm{f} = X \mathrm{f}$ so in this case it would be $X^\beta \partial_\beta \langle \alpha , Y \rangle$ like that, but this function $\langle \alpha , Y \rangle$ in comp notation is just $\alpha_\sigma Y^\sigma$ so this is the function, the partial derivatives of that:

$$X^\beta \partial_\beta (\alpha_\sigma Y^\sigma) = X^\beta (\alpha_\sigma Y^\sigma)_{|\beta} = X^\beta (\alpha_{\sigma|\beta} Y^\sigma + \alpha_\sigma Y^\sigma{}_{|\beta}) \tag{14}$$

Recall α is a function of space-time and Y is a function of space-time. We use (14) to substitute and put right in (13) so let me do that so it was:

$$\langle \nabla_X \alpha , Y \rangle = X^\beta (\alpha_{\sigma|\beta} Y^\sigma + \alpha_\sigma Y^\sigma{}_{|\beta}) - \langle \alpha , \nabla_X Y \rangle \tag{15}$$

Then we have to figure out this term $\langle \alpha , \nabla_X Y \rangle$ we know that this second part here is just:

$$\nabla_X Y = X^\delta (Y^\mu{}_{|\delta} + \Gamma^\mu_{\sigma\delta} Y^\sigma) \tag{16}$$

That's the literal translation of this expression into comp notation but then we're taking the dual space mapping with α so that we'll have so the only thing that it's left after this contraction is μ so what you're left with is you have to have an α_μ like that so it's going to be:

$$\langle \nabla_X \alpha , Y \rangle = X^\beta (\alpha_{\sigma|\beta} Y^\sigma + \alpha_\sigma Y^\sigma{}_{|\beta}) - \alpha_\mu X^\delta (Y^\mu{}_{|\delta} + \Gamma^\mu_{\sigma\delta} Y^\sigma) \tag{17}$$

What are we gonna do about this left hand side in (17) in comp notation? We're gonna have to break this down in the same way as we did with vectors, we're going to say this is:

$$\langle \nabla_X \alpha , Y \rangle = X^\beta \alpha_{\sigma\|\beta} Y^\sigma \tag{18}$$

We're trying to solve for $\alpha_{\sigma\|\beta}$, we want to know how to write that in component notation. Let's look over here (the right hand side of (17)), what I notice is if I just changed β to δ and σ to μ then terms would be identical and they're coming in at at different signs so they cancel each other:

$$X^{\delta}\alpha_{\mu\|\delta}Y^{\mu}=X^{\delta}\alpha_{\mu|\delta}Y^{\mu}-\alpha_{\mu}X^{\delta}\Gamma^{\mu}_{\sigma\delta}Y^{\sigma} \tag{19}$$

This has got to always be true for any vectors Y and X , you're gonna have to get this to work:

$$X^{\delta}Y^{\mu}[\alpha_{\mu\|\delta}]=X^{\delta}Y^{\mu}[\alpha_{\mu|\delta}-\alpha_{\sigma}\Gamma^{\sigma}_{\mu\delta}] \tag{20}$$

This is totally fair because for dummy indices, arbitrarily renamed right. This expression is supposed to be true for all vectors X and Y which allows me to write:

$$\alpha_{\mu\|\delta}=\alpha_{\mu|\delta}-\Gamma^{\sigma}_{\mu\delta}\alpha_{\sigma} \tag{21}$$

This is your expression for the comp covariant comp notation for the Covariant derivative of a co-vector. It's very similar to the vector, it's got the minus sign here and of course you're doing the contraction up down that way instead of up down the other way because that's the only other alternative

Now we understand the idea of a Covariant derivative of a vector $Y^{\mu}{}_{\|\delta}$, the Covariant derivative of a 1-form $\alpha_{\mu\|\delta}$. We understand how these expressions connect with the ordinary partial derivative $Y^{\mu}{}_{|\delta}$ and $\alpha_{\mu|\delta}$ and the connection on the manifold, we understand how these are driven right into that and ∇Y and $\nabla\alpha$ is the same thing just sort of jammed into this CFREE notation but it's really intended to mean the same thing and so now that we can do it for a vector and for a co-vector now the question is can we do it for an arbitrary tensor $\boldsymbol{Q}$ where $\boldsymbol{Q}$ in the comp notation would be $\boldsymbol{Q}^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q}$ so it would be a (p,q) rank tensor and showing that will be our next lesson.