Lesson 12: Parallel transport using operators

We're going to continue our discussion of Parallel transport not because we haven't done it enough we probably have done enough but because I'm not finished I think there's a lot more to look at we I'd like to do the whole idea of Parallel transport in CFREE notation because I'd like to have everybody develop a facility converting between CFREE and comp notation, I think that's an important skill to master for this understanding and learning and reading different books plus doing it just gives us another swing at understanding the concept of Parallel transport so there's really no harm in doing it again. We're not on a clock here and this is not a semester course where there are tests and there's a completion date so going through the same material in different ways that's not a problem to me.

I want to point out a little bit of an interesting thing about notation, we always have a curve when we're talking about Parallel transport we have a curve and we have a point on that curve and in that point on that curve we have a tangent space and that tangent space is a complete vector space and the basis of that tangent space and the coordinate basis would always be written as these partial derivatives ∂_μ , if it was the cotangent space the coordinate basis would be these 1-forms dx^μ but we're not dealing with the cotangent space, it's just the tangent space and all the vectors can be expressed in the component notation as $x^\mu \partial_\mu$, in the CFREE notation we don't bother with that and we just call all of the vectors inside this tangent space we just give them their name $\vec{X}$ and if it was in print, it would be in boldface but I might throw a vector symbol on top just because I have trouble with boldface on this whiteboard.

If I just pick one point on this curve and if the curve is parametrized by τ and maybe I start τ at zero and just go up to some higher values, there is really no limit to where τ can go, zero is just an arbitrary spot. I have to give the curve a name and the curve's name is γ so $\gamma(\tau)$ is the map from $\mathbb{R}$ to S and this point here that we're dealing with this point P , that's going to be $\gamma(\tau=0)$ or $\gamma(0)$. All I have is a vector sitting at that tangent space I've identified some vector though and my question is if I went to some other point $\gamma(t)$, this point has its own tangent space $T_{\gamma(t)}S$ this is the tangent space of $\gamma(t)$ whereas $T_{\gamma(0)}S$ is the tangent space at $\gamma(0)$. My question now is what vector here living in this tangent space $T_{\gamma(t)}S$ is parallel to $\vec{X}$? There's some vector in $T_{\gamma(t)}S$ that is parallel to this vector $\vec{X}$ which lives in $T_{\gamma(0)}S$. Notice I'm not talking about a Vector field right now, I'm saying I pick one vector at a point $\gamma(0)$ and I'm asking for its parallel counterpart at some other point $\gamma(t)$ and the reason I don't need a Vector field to talk about that is that parallelism, we now know, is defined by the metric so presumably our space S is … I'm sorry not the metric a connection (S,∇) and a connection, we haven't introduced the metric yet, there's no metric here at all but the connection is all I need to know to Parallel transport a vector from this point $\gamma(0)$ to this point $\gamma(t)$, I don't need the metric.

Eventually we'll talk about the metric and, as I've mentioned before, the metric defines a connection that will do just fine to determine parallelism and that is the foundation of how it's used in General Relativity but in the meantime until we have a metric as long as we have a connection on the space-time of some sort we can execute this Parallel transport so you can imagine doing this for all these different values of t , you know, t starts at 0 and t moves all the way up to unlimited values. $\gamma(t)$ moves along this curve and every single one of those points every single place you're going to have some vector that's parallel to $\vec{X}$. One question is how do we symbolize that vector and all of these different places and what I will suggest is that well we know how to find this vector we know about the Parallel transport equation so we know that that the vector at all of these different spots is going to satisfy Parallel transport equation, if we solve this equation:

$$\frac{\partial X^{\alpha}}{\partial \tau}+\Gamma^{\alpha}_{\beta\sigma} X^{\beta} \frac{d\gamma^{\sigma}}{d\tau}=0 \tag{1}$$

If we solved that we are going to end up with $X^{\alpha}(\tau)$. The solution of this these four differential equations that are coupled together will give us $X^{\alpha}(\tau)$ and what's nice is that this will be auto parallel on γ , that means I put in τ and I get the vector that is parallel to some initial condition so I have to solve this with the initial condition of $X(\gamma(0))$, that's my initial condition, that's this value right here $\vec{X}$ and once I've put in that initial condition the Vector field I get as an answer to these differential equations, that Vector field will be auto parallel and therefore this will now be $\vec{X}(s)$ if the point is $\gamma(s)$ and then this is $\gamma(r)$ and then this would be $\vec{X}(r)$ and then this would finally be $\vec{X}(t)$ and that would be an auto parallel field. We have to use the comp notation to develop the differential equation to actually do the solution but in the meantime if we want to stay in the CFREE notation, we presume that this is possible and we just define this Auto parallel field $\vec{X}(\tau)$ and we say that this is $\vec{X}(0)$, $\vec{X}(s)$, $\vec{X}(r)$ and $\vec{X}(t)$ and they're all parallel to one another because we derive this expression for $\vec{X}$ from the Parallel transport equation (1) so with this notation in hand we are going to generate an expression for the Parallel transport equation using the comp notation from the beginning and the way we do this is we're going to introduce an operator Ω and we'll call this operator $\Omega_{t,s}$. All operators have to be defined as mappings and this operator will be:

$$\Omega_{t,s}: T_{\gamma(s)}S \rightarrow T_{\gamma(t)}S \tag{2}$$

At first it seems kind of ridiculous but the way the operator works is:

$$\Omega_{t,s}\vec{X}(s)=\vec{X}(t) \tag{3}$$

Remember that $\vec{X}(\tau)$ is an auto parallel field or a Vector field that is auto parallel on the curve γ so it's only really defined on γ and γ is a curve that is parameterized by τ . Notice by the way you can name these curves in a lot of different ways, well in at least two different ways that we're dealing with here. τ is the parameter for the curve and the name of the curve is γ because γ is the map that takes τ which is a real number and gives you a curve in S so you can name the curve after γ but it's not like τ is the parameter for other curves, it's not like there's another $\varphi(\tau)$ and you have to figure out which curve you're going to. Each curve is gonna have its own parameter and because of that τ could name the curve too. τ could easily be the name of the curve so saying $\vec{X}(\gamma(\tau))$ is redundant because τ and γ both really identify the curve so you can make the mapping γ kind of implicit and just say $\vec{X}(\tau)$ understanding that τ is the parameter for some curve and that curve may even have a name γ but I've seen some treatments that just leave the name out altogether.

What we've done, it seems almost silly, it's like we've made a complicated version of $\vec{X}(t)$ but this operator is the Parallel transport operator: when applied to $\vec{X}(s)$ it will give you $\vec{X}(t)$ where we understand that we're really talking about $\vec{X}(\tau)$, is the function we're talking about which is really a Vector fields along a curve $\vec{X}(\gamma(\tau))$. This is all sort of introducing certain forms of shorthand. Now within this notation we have a way of expressing the Parallel transport equation:

$$\nabla_{\dot{\gamma}} \vec{Y} = 0 \tag{4}$$

That's the Parallel transport equation. In our new notation the way we're going to understand that is:

$$\left. \frac{d}{ds} \right|_{s=t} \Omega_{t,s} \vec{Y}(s) = 0 \tag{5}$$

This is the Parallel transport equation in the abstract notation that utilizes this method (3) of Parallel transport. I'm trying to understand what this is, this is the derivative with respect to s, $\vec{Y}$ is an auto parallel field and s is the parameter value that defines the place on the curve, this could be $\vec{Y}(\gamma(s))$ and this operator $\Omega_{t,s}$ is also a function of s because you have to choose s and t so (5) is the rate of change of this object with respect to s and it's evaluated when $s=t$ which is the identity by the way because $\Omega_{t,t}$ is going to be the identity. Moreover:

$$\Omega_{r,t} \circ \Omega_{t,s} = \Omega_{r,s} \tag{6}$$

There's that transitiveness, that's because along this curve γ, it's unique, there's only one parallel vector to this initial vector $\vec{X}(0)$ at this point $\gamma(0)$ and if you Parallel transport that to $\gamma(r)$ there's only one parallel vector in $\gamma(r)$ that's going to be parallel to $\vec{X}(0)$ so to go from $\gamma(0)$ to $\gamma(r)$ you've got to go through the same batch of parallel vectors, it's because the differential equation has a unique solution and because of that that ultimately translates into this statement here (6). These are by the way are finite distances where finite means they're finite values of the parameter τ so this isn't an infinitesimal distance like we've talked about previously this is these are finite distances of τ and the word distance is a little awkward because we don't have a metric so what I should say is finite changes in the value of the real number τ which parameterize the curve.

In some sense we do have a metric, if you have a curve and τ is a parameter on the curve you know in some sense τ is a bit of a metric it does define the distance along the curve in terms of some parameter τ the problem is you can easily re-parametrize γ without any trouble that would be similar to just changing the scale of the curve or something like that. This expression (6) comes from the uniqueness of the solution of this differential equation (1). Now I want to show why this expression (4) and this expression (5) are the same and to do that we're going to start, we've got to remind ourselves let's remind yourselves that:

$$Y(t) = \Omega_{t,s} Y(s) \tag{7}$$

The auto parallel vector at t is the auto parallel Vector field at s run through this operator, well we can think of this in terms of components and you know typically to prove these things for our CFREE notation we always dive into component notation to do it but this vector $[Y(t)]^{\nu}$ has components so let's think about the different components, well that would be the components of this guy $[\Omega_{t,s} Y(s)]^{\nu}$ because they're the same and then you've got to look at this and realize $Y(t)$ is a vector so this vector can be $Y^{\alpha} \partial_{\alpha}$ and this guy here it's also a vector and each of these components here is some kind of collection of these components there so this can be understood as ω operating as a matrix we can think of $[\Omega_{t,s}]$ it's going to have a ν component and new component σ,

$$Y^{\alpha}\partial_{\alpha}=[\Omega_{t,s}]^{\nu}_{\sigma}\,Y^{\sigma}(s) \tag{8}$$

If I erase this expression here (left hand side) is sort of the full component version of this expression here (7), I guess I have to include that in front on this side and what you can see is that this is just a summation over the components of $Y(s)$:

$$[Y(t)]^{\nu}=[\Omega_{t,s}]^{\nu}_{\sigma}\,Y^{\sigma}(s) \tag{9}$$

Each of the components of $Y(s)$ is going to contribute somehow to the new component of the auto parallel version at t and that amount of change is going to be all wrapped up and expressed in the components of this somewhat mysterious operator that moves $Y(s)$ to $Y(t)$. Now when I say it's a mysterious operator remember, the nature of that operator simply is tell me where we start tell me where we end and I'll solve this equation use the initial condition of where we start and plug in the value of the parameter for where we end and I'll change those vector components into those, so it's not a mysterious operator it's really describing that process and it turns out that that process, these are going to be complex functions of s and t and in particular since it's a function of s taking its derivative with respect to s we're going to ultimately see that is going to define the equation for Parallel transport. We have this expression (9) for the components of $Y(t)$ and just to be completely clear about it we can write down like this, we can say:

$$[\Omega_{t,s}Y(s)]^{2}=[\Omega_{t,s}]^{2}_{0}Y^{0}(s)+[\Omega_{t,s}]^{2}_{1}Y^{1}(s)+[\Omega_{t,s}]^{2}_{2}Y^{2}(s)+[\Omega_{t,s}]^{2}_{3}Y^{3}(s) \tag{10}$$

Now we know that $Y(t)$ satisfies the Parallel transport equation so we know that that satisfies (1):

$$\frac{dY^{\mu}(t)}{dt}+\Gamma^{\mu}_{\alpha\beta}\frac{d\gamma^{\alpha}(t)}{dt}Y^{\beta}(t)=0 \tag{11}$$

Remember, when we write down Y^{β} , we're thinking Y^{β} is a function of t . As a reminder whenever you look at a differential equation understand what are you solving for here? We are solving for a function $Y(t)$ which it's gonna solve this whole differential equation and this is a known quantity (the term with γ) because we presume that we know the path, we know this function γ , we know what path were interested in moving along and finding the Parallel transport along and the connection is part of the space-time, it's a given piece of the space-time (S,∇) , we have a connection.

I've converted everything back into comp notation, it's just this thing we do when we do these proofs, it's sometimes easier to do the proofs in comp notation and now I'm trying to prove that this expression here (4) is equivalent to this expression here (5) and I'm going to do it stepping through comp notation. Now so I can rewrite this expression as:

$$\frac{d[\Omega_{t,s}]^{\mu}_{\sigma}\,Y^{\sigma}(s)}{dt}+\Gamma^{\mu}_{\alpha\beta}\frac{d\gamma^{\alpha}(t)}{dt}[\Omega_{t,s}]^{\beta}_{\sigma}\,Y^{\sigma}(s)=0 \tag{12}$$

I've now changed it to a function of s using (9). Remember the connection is a function of t too, it's just that there's no derivatives being taken on the connection.

Then what we note is we note that we have $Y(s)$ in two places (12) and therefore we can pull it out and the fact that $Y(s)$ is arbitrary will allow us to make this 0 so:

$$\left[\frac{d[\Omega_{t,s}]^{\mu}_{\sigma}}{dt}+\Gamma^{\mu}_{\alpha\beta}\frac{d\gamma^{\alpha}(t)}{dt}[\Omega_{t,s}]^{\beta}_{\sigma}\right]Y^{\sigma}(s)=0 \tag{13}$$

The reason you can get rid of $Y(s)$ is because $Y(s)$ is arbitrary, we're allowed to have any Vector field to start with, there's no conditionals on this so in order for this whole thing to equal 0 if one factor could literally be anything the only way that can always be true is if this thing in brackets was always 0 so that's why we allow ourselves to then get rid of $Y(s)$ and just write the expression this way:

$$\frac{d[\Omega_{t,s}]^{\mu}_{\sigma}}{dt}+\Gamma^{\mu}_{\alpha\beta}\frac{d\gamma^{\alpha}(t)}{dt}[\Omega_{t,s}]^{\beta}_{\sigma}=0 \tag{14}$$

Once we've done that, that's our expression (14). I've done all that just for that last line:

$$\frac{d[\Omega_{t,s}]^{\mu}_{\sigma}}{dt}=-\Gamma^{\mu}_{\alpha\beta}\dot{\gamma}^{\alpha}[\Omega_{t,s}]^{\beta}_{\sigma} \tag{15}$$

That's the statement I wanted to demonstrate and here I'm using $\dot{\gamma}$ using this for the derivative. We're already at something that's beginning to look like what we want to show, we wanted to show (5) and I already have something now that expresses a derivative with respect to one of these sub-scripted variables. I have an expression for that now in terms of the object itself, the connection coefficients and the derivative of the curve and that's very useful. I can go even further if I evaluate this thing at $t=s$, if I do that I end up with an expression that looks like it looks like this:

$$\left.\frac{d[\Omega_{s,s}]^{\mu}_{\sigma}}{dt}\right|_{s=t}=-\Gamma^{\mu}_{\alpha\beta}\dot{\gamma}^{\alpha}\delta^{\beta}_{\sigma}=-\Gamma^{\mu}_{\alpha\sigma}\frac{d\gamma^{\alpha}}{dt} \tag{16}$$

That expression is sort of the final answer when you evaluate (15) at $s=t$ which is what we are going to do, eventually that's part of our ultimate goal. We've kind of gotten this in place ready to go meet the next substitution and so now we just can write down the whole thing and we would write it the following way we will write:

$$\left.\frac{d[\Omega_{t,s}Y(s)]^{\mu}}{ds}\right|_{s=t}=\left.\frac{d[\Omega^{-1}_{s,t}Y(s)]^{\mu}}{ds}\right|_{s=t}=\left.\frac{d}{ds}[\Omega^{-1}_{s,t}]^{\mu}_{\beta}\,Y^{\beta}(s)\right|_{s=t}+\left.\Omega^{-1}_{s,t}\,\frac{d}{ds}Y^{\mu}(s)\right|_{s=t} \tag{17}$$

That's the definition we're trying to prove, we're trying to show that this object here is our left-hand side of the Parallel transport equation (1) and I'm going to just turn it into that using everything we've just learned so the way I'll do that it's actually kind of tricky.

I will go to the inverse and I'm going to swap the subscripts s and t. This is the same as this because when I switch from s to t is the same as taking the inverse so if I switch from s to t and then take the inverse is exactly the same. Then I used the product rule. Then I make the one little tricky maneuver, that is, I take this guy (first term) and I replace it with:

$$-\frac{d}{ds}[\Omega_{s,t}]^{\mu}_{\beta} Y^{\beta}\Big|_{s=t} + \frac{d}{ds} Y^{\mu}(s) \tag{18}$$

I'll show you why we can do that in a moment. Let's have a look over here at the change from $\Omega^{-1}_{s,t}$ to losing the inverse and just putting a negative sign up there. Let's see if we can understand that. To understand that let's see if we can break it down. I'm gonna break it down to little segment of our curve this is the point represented by the parameter value s and this is the point represented by the parameter value t and this is the auto parallel field at s and then this is the auto parallel field that t. We know (9) so when we look at this expression here we look at this expression here from (17):

$$\left(\frac{d}{ds}[\Omega^{-1}_{s,t}]^{\mu}_{\beta}\right) Y^{\beta}(s) \tag{19}$$

We have this expression so our question is how do we interpret this expression? First of all what's important to understand is that there's some parentheses here, because we developed it from a product rule so it's this first derivative part, so now this is the change, it's a vector change, these are the components of a vector at s so you would have the vector at s the operator that goes from s to t and then the vector of t. In (19) you have the vector at s and this is an operator that goes from s to t because it's the inverse of an operator that goes from t to s so this thing here is going to ultimately represent the change in the vector components of what the components were of $Y(s)$ to what the new parallel components will be at t so if we just move the components of s over here we might get something that would look like that, say, and that would be this vector here sitting at t has the exact same components at s which would be what Parallel transport was in the absence of any connection, if the connection was zero and it was flat space you would just take the components here and move them here and call it parallel but that's not the reality.

If I did that and moved it here I would get this vector but the real parallel vector is this one which has been moved or shifted, its components are different because of the connection and that difference is this vector here the vector between what would have been had it been flat space and what it is because it's curved space and that's kind of what this thing is telling us (19), this thing is part of that change, it is where the other part is:

$$[\Omega^{-1}_{st}]^{\mu}_{\beta} \frac{d}{ds} Y^{\beta}(s) \tag{20}$$

This is part of what that change is. The point is now if we took the same logic and we dragged $Y(t)$ and just put its components in s we would end up with something like this which is not the Parallel transport of $Y(t)$ to s what that would the parallel of $Y(t)$ to s should be this guy $Y(s)$ by the way we've constructed everything. The difference there is this vector here but you can see these two things will have to be opposite of each other. If I do the inverse Parallel transport here this vector which corrects for the non flatness is going to just be the opposite of if I did the Parallel transport here so that's what allows me to eliminate this inverse by just changing the sign of all of these little vector components and now I've already written down the next step after we take this guy here (18) and we are going to now substitute in for this expression based on what we learned here (16). I can make this substitution (18) right here:

$$\Gamma^{\mu}_{\alpha\beta}\frac{d}{ds}\gamma^{\alpha}Y^{\beta}(s)+\frac{d}{ds}Y^{\mu}(s)=\nabla_{\dot{\gamma}}Y \tag{21}$$

That comes from that prior work we did and then the $Y^{\beta}(s)$ comes down and then I evaluated at $s=t$ and then add this (second) term into it and I get this term and then when I look at this left hand side of (21) I see the familiar expression which I guess, if I reorganize it, I would get:

$$\frac{dY^{\mu}}{ds}+\Gamma^{\mu}_{\alpha\beta}\frac{d\gamma^{\alpha}(s)}{ds}Y^{\beta}(s)=\nabla_{\dot{\gamma}}Y \tag{22}$$

We know this is the comp form of this expression $\nabla_{\dot{\gamma}}Y$ and then ultimately that is what we were trying to show, we were trying to show that:

$$\nabla_{\dot{\gamma}}\vec{Y}=\frac{d}{ds}[\Omega_{ts}\vec{Y}]\Big|_{s=t}=0 \tag{23}$$

That's what we wanted to show and we did, we started up there in the far corner here (17) with this expression and we worked our way through it and we ended up with this expression (22). Now we have using the operator notation we have way of writing the Parallel transport equation which is this (23), and that's a worthy goal for one lesson just to see yet a different form of notation architecture at work this operator notation and you should be relatively comfortable moving between them but the one you really need to be comfortable with to really get at the introductory and intermediate General Relativity is of course the comp notation, this stuff here is really important and from time to time I've seen older books to do proofs in this form but it really is all the same mathematics I don't want to say it's just a matter of notation because the insights from here to here they're they're very different this is very computational and we actually do a lot of proofs, you even saw in order to make a proof about this abstract notation we immediately went into components so you know it's very tempting to just turn everything into components work with components and then sort of turn them back because we know this definition from our previous work.

We know this definition (23) from this work today, so the next lesson what we'll talk about is we'll talk a little bit about parallelism of general tensors there's a nice abstract proof of how to define parallelism for general tensors and we'll use both of these notation systems to do that and then we'll see how that links up with the definition of covariant derivatives for tensors of arbitrary rank and from there we will go into defining a very an abstract approach of defining the directional or partial derivative and then that will lead us into the notion of the metric connection where we are basically claiming that Parallel transport should not change the inner product between vectors so if I take two vectors in one tangent space and find the metric between them or that if I compare them using the metric then I should be able to Parallel transport those two vectors to any other point and compare it with the metric there and I should get the same answer and that's what ties the metric to the connection. We'll see you next time.