Lesson 8: Intro to the metric connection and the induced metric

We're going to begin a series of topics where we're going to wander through the subjects that will get us to the notion of the connection. We've already talked a lot about connections so I'm not going to be teaching what the connection is from the start but what I'm going to do is I am going to describe the connection that's used in General Relativity we've talked about the connection in a mathematical sense and in the sense of Differential geometry and some of our work with tensors and manifolds but what about the connection as far as General Relativity goes, how does that apply. In General Relativity we have our basic model of space-time which is a manifold that has coordinates and we've already explained that at each point in space-time we have Tangent spaces and these tangent spaces contain essentially, well the tangent space is a Vector space but associated with this is a covector space and all the equivalent Tensor products spaces we still got that going and we also know that in order to make this space-time if we call the space-time S and $S=\{\text{Points in Space-time}\}$ and it's a manifold so this is a manifold because it's a manifold we know there's a coordinate system (x^0, x^1, x^2, x^3) and all of those points make up the space-time but even as a manifold and even with a coordinate system where we can make patches and lay down coordinates using the charts you could say that S as a set of events in space-time, actually a set of points M and an Atlas A i.e. $S=(M, A)$ that's what that's all about as the atlas is what gives it the coordinates in the Atlas A is itself a set, a set of charts $A=\{(U_\alpha, \gamma_\alpha)\}$ and it's the set of all those charts which map every point into $\mathbb{R}$ for in our case into $\mathbb{R}^4$ and so the space-time is that manifold but it's more, the space-time is actually the manifold and a metric, the metric is what makes the difference here because the metric is what binds together all of the inner products and all of these points a metric defines an inner product at every every point in space-time which allows you to do things like the (v, v) , you take two vectors from the underlying tangent space and determine something like a magnitude, it allows you to do all of this measurement so that's cool.

We know that that's how our model of space-time is going to work and ultimately we know that the Einstein equation is going to take matter, stick it in here $\boldsymbol{G}_{\mu\nu}=8\pi \boldsymbol{T}_{\mu\nu}$ and then space-time is gonna pop out here $g_{\mu\nu}$ and space-time means the metric so that's how you're gonna do it, you can start with matter which is all forms of energy in the right side of the equation and the left side of the equation is gonna give you your metric and then these two things together the manifold and the metric that's the arena of General Relativity and that's what the subject really is all about.

You may remember when we formulate the laws of physics we do it with usually differential equations and we've already described in previous lessons and on “What is a tensor series” about the connection and how that's related to the Covariant derivative so if I have a tensor $\boldsymbol{T}_{\alpha\beta}$, I can take its Covariant derivative which frequently is denoted with a semicolon $\boldsymbol{T}_{\alpha\beta;\gamma}$ with respect to space-time indices or I can take a better notation $\boldsymbol{T}_{\alpha\beta\|\gamma}$, the notation I'd like us to use with these two lines.

We talked about that but that in order to get through these two Covariant derivatives we need to create the connection and here's an example of a connection was Γ and the connection and the metric are not the same thing, the metric just tells us how to take an inner product at any given point, the connection is what tells us how parallelism works how can we understand a vector in the tangent space at some point and relate it with a vector that's in the tangent space in some other point. The point is the connection breaks the seal I mean even though you have a metric that tells you what the inner product is at every point in space-time that doesn't do anything to connect this tensor product space in one point with that tensor product space in some other point. Those tensor product spaces or those vector space this tangent spaces are still an infinite universe apart, they're totally different objects even though they have a metric which tells me they're two separate inner products.

For example I could have $g_{\alpha\beta}(P)$ and when I write this what I'm really saying is $g_{\alpha\beta}(x^{\mu}(P))$, that's how the manifold works I can take the point there I get the coordinates there and $g_{\alpha\beta}$ is a function of the coordinates so here I have $g_{\alpha\beta}(x^{\mu}(Q))$. I have a $(0,2)$ Tensor field called the metric and it's a continuous function the coordinates it's a beautiful smooth function of coordinates and it tells me the inner product, there and likewise this will tell me the inner product there but that does not establish any kind of relationship between this tensor product space and that tensor product space. in order to do that I need a connection so you can now argue that for a space-time I need a manifold, I need a metric, in order to do General Relativity and I need a connection, I need 3 things $(S, g_{\alpha\beta}, \Gamma^{\gamma}_{\mu\nu})$. What we're now going to start working towards the direction of going back to this catalog of space-time which is sort of our reference book here and we're looking here at **1.3 Basic objects of the metric** in the Catalog of space-times and the first thing in this list (1.3.1) is this object here which is called the Christoffel symbol of the first kind and (1.3.3) is the Christoffel symbol of the second time let's look at that one because that one has an object on the left that looks like the connection in fact it gives the connection and notice how it takes that connection and it defines the connection in terms of the metric, the inverse metric, well the inverse metric is just defined in the obvious way $g^{\mu\nu} g_{\mu\nu} = 1$ but more importantly it takes the derivative of the metric with respect to space-time where these symbols are:

$$\Gamma^{\mu}_{\nu\lambda} = \frac{1}{2} g^{\mu\rho}\left(g_{\rho\nu,\lambda} + g_{\rho\lambda,\nu} - g_{\nu\lambda,\rho}\right) \text{ where } g_{\rho\nu,\lambda} \stackrel{\text{def}}{=} \frac{\partial g_{\rho\nu}}{\partial x^{\lambda}} \qquad (1)$$

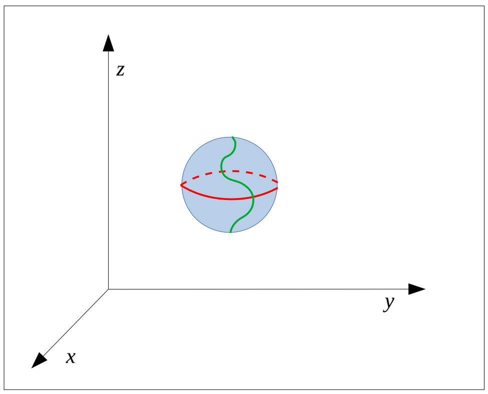
This is the connection, is given in terms of the metric so now if we go back to our little thing here our space-time if that's what we believe then the metric is all we need $(S, g_{\alpha\beta})$ so now what we have to understand is why is it that if we have a space-time, if we have a metric how does that give us a connection? What's this what's the relationship what's the connection between the metric and the connection and this just goes to show the last step is everything in this level of our understanding. It's all about the metric, no metric, no nothing is what some famous astrophysicist once said. With the metric you know how far apart they are the fact that the coordinates are adjacent to each other or the coordinates have similar values that doesn't mean anything unless you have a metric and in fact we already know that it's some bit screwy because we already know that in Special Relativity I could have a point here and a point here and they are actually 0 apart if they're on a line of 45º in a (t, r) diagram that sphere here and that sphere there 0 apart if they're connected by a 45º line because of the Lorentz nature of the metric but ultimately those distances, you got to understand them but there is a way of taking this measurement of distances and using that to tie together the tangent spaces and then understand parallelism. Remember that's what this was all about taking a vector that was established at one point P and a vector that's established at a point Q and saying there's a vector here in this tangent space, the tangent space of Q of the space-time and there's a vector here at the tangent space of P of the space-time and this vector here is parallel to that vector there we can establish parallelism if we have a connection and now we want to ultimately learn as the catalog shows that that connection is related to the metric and we need to understand why that is. That's our goal for the next lecture or two or three, it's actually a beautiful complicated subject and its really worth digging into because there's a lot of pieces to it.

Before we begin I want to have a good understanding of a metric a little bit better, I want to be able to understand the classical nature of the metric, when I say classical I mean classical Differential geometry tells us something about the metric and that tells us a little bit about how we need to think of the metric in terms of General Relativity. Let's begin with some an elementary discussion of the metric that is relevant to what we've done before it's going to connect to the conformal coordinates we've done before so let's begin with that. Something that's easy for us to understand is Euclidean space with Cartesian coordinates. We know that the metric for Euclidean space is gonna always be written as:

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2}$$

I'm using this sort of the casual notation of a metric by describing the line element so here I would say:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$



We know that that's the metric of our standard geometry of sort of classical Euclidean geometry so the question we want to understand is I want to kind of exercise the nature of the metric in a very classical sense just to warm us up and that's gonna take this full lesson. What I'm going to imagine a surface in this 3D space and I'll just choose a sphere, spheres are nice classical simple surface it's often used for this example and that sphere is a set of points such that:

$$x^2 + y^2 + z^2 = R^2 \text{ where } R > 0 \text{ is the radius of the sphere} \tag{4}$$

That's a surface so we understand that this is a 2D surface in or embedded in a 3D manifold. Now we know the metric of the underlying 3D manifold so now we talk about a curve just in the surface for example, well you could talk about this curve here (in red) say mostly around the circle, well we could talk about some arbitrary curve on the surface some curve that connects one pole to the other (in green) but it doesn't have any particular symmetry on the surface it is fixed and we can calculate the length of those two curves and they are curves, so given the fact that they're curves what do we remember about curves? We remember that curves we describe it with a parameter say the parameter is $0 < \tau < 1$ and then the curve itself could be, say $\gamma(\tau)$ and it maps onto the sphere. You can think of it as mapping on to the sphere or into the 3D space because this curve is in the 3D space is just on this sphere so you can think of it both ways you're either mapping onto the surface of the manifold or onto the 2D. The sphere is a 2D manifold embedded in a 3D manifold. The 3D manifold only needs one chart, the 2D manifold would need two charts, we discussed that in the “What is manifold” lecture series so you can think of this as going both ways:

$$\gamma(\tau) = \left(\gamma^1(\tau), \gamma^2(\tau), \gamma^3(\tau)\right) = \left(x(\tau), y(\tau), z(\tau)\right) \tag{5}$$

In our standard notation it would be $\left(x^1(\tau), x^2(\tau), x^3(\tau)\right)$. Remember this is a manifold and as a 2D surface it has two coordinates and those coordinates could be (θ, φ) m those are the standard angular coordinates in the surface of a sphere as long as this is all in one chart and then you could instead say consider:

$$\gamma(\tau)=\left(\gamma^{1'}(\tau),\gamma^{2'}(\tau)\right)=\left(\theta(\tau),\varphi(\tau)\right) \tag{6}$$

I could have written this as $\left(x(\tau),y(\tau),z(\tau)\right)$ and that would have been $\left(x^{1'}(\tau),x^{2'}(\tau)\right)$, I'm just using a few different notations here, There's two ways of doing it and likewise I can measure the length of that curve in both coordinate systems and to measure that length I'll need the metric so I could use the metric, let's call the metric of the Euclidean space G_E or if I'm using this set of coordinates (θ,φ) I could measure that curve using the metric of the sphere which would be G_S . What's important in what I want to show today is that G_E is an easy metric but G_S not so much, I mean we know what it is but in principle we should be able to take an arbitrary surface, not just a nice sphere we know what it is because we know all about spheres but it could be any surface. It seems like if you know G_E you should be able to figure out what's G_S , if I know the metric of the underlying space in which this manifold is embedded that's G_E shouldn't I be able to figure out what's G_S ? The answer is yes and we are going to do it but we're gonna do it for a case that's a little more complicated than this, we're gonna use the 4D sphere so we're gonna exercise thinking in 4D and we're gonna exercise this thing called the Induced metric which is the metric of a sub manifold of a manifold when you know the metric of the embedding manifold and that's what we're gonna do it now.

We can begin this by reminding ourselves about the basic sphere and then we're going to generalize it to one the extra dimensional play with that for our it's going to be a sort of an entertaining arena to test our the concept of the induced metric but this sphere is going to be the set of these points that have (4) and that's in 3D so this is the set of all points, all triplets (x,y,z) that satisfy(4), that's the sphere and then on top of that we want to create the metric of the sphere G_S and the metric of the sphere we want to get that from the metric of Euclidean space but we also want to play more with the more interesting thing which is I want to talk about the 4D sphere which I don't know quite how to draw it I know I mean we don't need to draw very much but if I were to draw it I kind of draw out the way people try to draw a Tesseract which is one sphere and then a little sphere sort of inside it. The definition is:

$$x^2+y^2+z^2+w^2=R^2 \text{ where } R>0 \text{ is the radius of the sphere} \tag{7}$$

It's the set of all quadruples (x,y,z,w) that satisfy this relationship (7) and it will also have a metric called the “Glome” metric G_{Gloam} because a 4D sphere is often called a “Glome” but that can be derived from G_{E4} . The “Glome” 3D space embedded in 4D space, it's embedded in $\mathbb{R}^4$ so the metric from $\mathbb{R}^4$ should be able to derive the “Glome” exactly the same way the metric from $\mathbb{R}^3$ derives the sphere and in fact this is the one we're gonna do, it's gonna take a long time but we're gonna do it.

All the points on the inner sphere have the same (x,y,z) coordinates as all the points in the outer sphere but they have different w coordinates. Don't worry so much about drawing it because you can't do it nobody can draw it is just not visualizable but it is mathematically very clear. We have the “Glome” and we have the sphere and I'm gonna try to talk a little bit about both in parallel because one is what's gonna give us intuition for the other although it's very pretty straightforward.

$$X^n=r\cos\lambda\prod_{k=1}^{n-2}\sin\varphi_k \tag{8}$$

Now we can consider r to be a coordinate variable. For the familiar case of 3D this gives the usual Spherical coordinates if we assign $X^1 \equiv z$, $X^2 \equiv y$, $X^3 \equiv x$, $\varphi_1 \equiv \theta$ and $\varphi_2 \equiv \varphi$:

$$\begin{cases} X^1 = r\cos\varphi_1 \\ X^2 = r\sin\lambda\sin\varphi_1 \\ X^3 = r\cos\lambda\sin\varphi_1 \end{cases} \text{ or } \begin{cases} z = r\cos\theta \\ y = r\sin\theta\sin\varphi \\ x = r\sin\theta\cos\varphi \end{cases} \tag{9}$$

The “Glome” is the 3-sphere of radius r and is analyzed using the 4D case:

$$\begin{cases} X^1 = r\cos\varphi_1 \\ X^2 = r\cos\varphi_2\sin\varphi_1 \\ X^3 = r\sin\lambda\sin\varphi_1\sin\varphi_2 \\ X^4 = r\cos\lambda\sin\varphi_1\sin\varphi_2 \end{cases} \tag{10}$$

This is the transformation from “Glome” coordinates $(r, \varphi_1, \varphi_2, \lambda)$ into Cartesian coordinates (X^1, X^2, X^3, X^4)

If I establish that the sphere has R equals some fixed value then this becomes

$$\begin{cases} z = R\cos\theta \\ y = R\sin\theta\sin\varphi \\ x = R\sin\theta\cos\varphi \end{cases} \tag{11}$$

This side is 2D. Now I have the 3D of the $\mathbb{R}^3$ space (x, y, z) in (9) and I have the surface of the sphere S in (11). The whole point being that I can use the coordinates on the surface of the sphere to calculate the coordinates in $\mathbb{R}^3$ space, that's what this coordinate transformation does. It only takes two coordinates to establish any point on the sphere because the sphere is a 2D surface which is embedded in a 3D space but those dimensions are not all completely independent they're bound by (4) so you lose one degree of freedom. The “Glome” is a 3-sphere and it has the same kind of thing it has the same kind of coordinate transformation and that's (10). Now we end up adding an angle variable θ . θ and φ are not the same, they're very different because $0 \le \theta \le \pi$ whereas $0 \le \varphi \le 2\pi$ and it turns out that when you go to higher dimensions and this is not just the 4th dimension, it could be any dimension, it doesn't really matter what you end up adding our additional variables similar to θ there's only one similar to φ so we start naming the θ type variables, it's conventional to switch them I call $\varphi_1, \varphi_2, \varphi_3$ etc for an n dimensional sphere those would all be angles that belong to $(0, \pi)$ but $0 \le \lambda \le 2\pi$.

I kind of changed the notation, the reason I'm doing it is (10) is standard for the paper I learned about n dimensional Spherical coordinates and (9) is standard you see in Physics books. If I set R equals to some constant then I have reduced this side to an object with 3 coordinates:

$$\begin{cases} X^1 = R\cos\varphi_1 \\ X^2 = R\cos\varphi_2 \sin\varphi_1 \\ X^3 = R\sin\lambda\sin\varphi_1\sin\varphi_2 \\ X^4 = R\cos\lambda\sin\varphi_1\sin\varphi_2 \end{cases} \tag{12}$$

This is a 3D object embedded in a 4D space and that is called the "Glome". It depends whether you're a topologist or geometer by the way, it's a 3-sphere in the sense its surface takes 3 coordinates or it's a 4-sphere because it lives in 4D space so it depends on whether you're a topologist or a geometer and I forget which one calls it which. Now the question we're going to ask is what is the metric of the "Glome"? We know what this is we know the coordinate transformation already now we're going to ask what is the metric for the "Glome", the induced metric from 4D $\mathbb{R}^4$? We know the metric on $\mathbb{R}^4$:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{13}$$

That's the standard Euclidean metric for $\mathbb{R}^4$, now we want to know what is the metric is on the surface of the "Glome" i.e. what is $g_{\varphi_1\varphi_1}, g_{\varphi_2\varphi_2}, g_{\lambda\lambda}$. Maybe there's a $g_{\varphi_1\varphi_2}$, there could be a cross term here, there could be a $g_{\lambda\varphi_1}$ for example so we want to do that calculation and that's the goal of the rest of this lecture. We will start by writing down the surface of the globe in a very classic vector notation if I were to write down $\vec{r}(\varphi_1, \varphi_2, \lambda)$. I want to know the point in 4D Euclidean space that is the same as given these 3 parameters, I want to know the point in 4D Euclidean space and I want to use sort of this classic vector notation, how would I do that and I would do that by just writing it out in this fashion I need a 4D sort of vector object here and then I would stick (12) in this stuff. I'm going to assume that $R=1$.

$$\vec{r}(\varphi_1, \varphi_2, \lambda) = \begin{pmatrix} R\cos\varphi_1 \\ R\cos\varphi_2 \sin\varphi_1 \\ R\sin\lambda\sin\varphi_1\sin\varphi_2 \\ R\cos\lambda\sin\varphi_1\sin\varphi_2 \end{pmatrix} \in \mathbb{R}^4 \tag{14}$$

Now we can actually think in terms of vectors, when I say in terms of vectors I mean it's in terms of little pointy things outside in Euclidean space this is all very basic geometry so we can lean on all of that classic stuff we learned before I told you to forget it all however this is still going to be interpreted as an element of $\mathbb{R}^4$. We've implicitly put in an R but I'm thinking that if I want to choose $R=1$ so I don't have to carry it around. This is a description of the "Glome" in 4D space and we could likewise have done the same thing for a sphere in 3D space:

$$\vec{r}(\varphi_1, \lambda) = \begin{pmatrix} R\cos\varphi \\ R\sin\lambda\sin\varphi_1 \\ R\cos\lambda\sin\varphi_1 \end{pmatrix} \in \mathbb{R}^3 \tag{15}$$

That (14) is our “Glome” surface so with that I can now map in space this “Glome” object. Now my question is: so I'm going to be dealing with things in the surface of this object and I'm looking for the induced metric so it stands to reason that I'd be interested in the tangents spaces of the “Glome” which if it was a regular sphere in 3D space I'd be curious about if I take one point I want to know about this tangent plane to that point, every point will have a tangent plane and I'd like to know about some basis vectors that cover that tangent plane in other words I want to create the basis vectors that point into this tangent plane and cover the whole plane and I want to know this in the underlying 4D space. I want to find the basis vectors in 4Dspace for the tangent plane of any point on the sphere, likewise for the “Glome” I want to find the same thing but each point has a tangent space a 3D tangent space and I want to know 3 vectors that cover that tangent space for every point on the “Glome” so how do we do this?

By analogy, we're going to do this for the 3D and for the 4D case, we're going to construct 3 vectors $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_{\lambda}$, these 3 vectors are going to be basis vectors for the tangent space at the point $(\varphi_1,\varphi_2,\lambda)$ and we're going to construct them from (14) and the definition for that construction is the same as it would have been for the sphere it's going to be:

$$\vec{e}_{\varphi_1}=\frac{\partial\vec{r}(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}=\begin{pmatrix}\frac{\partial x(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\\ \frac{\partial y(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\\ \frac{\partial z(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\\ \frac{\partial w(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\end{pmatrix} \tag{16}$$

I think it's worthwhile to drag $(\varphi_1,\varphi_2,\lambda)$ around, in any reasonable textbook it will be shown without it because they would assume that you know but I actually I think it's worth dragging it around but now to make connection with what we have done this could also be better interpreted and probably is more correct well it's a little more error-free if you write this as:

$$\vec{e}_{\varphi_1}=\left(\frac{\partial x(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\right)\partial_x+\left(\frac{\partial y(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\right)\partial_y+\left(\frac{\partial z(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\right)\partial_z+\left(\frac{\partial w(\varphi_1,\varphi_2,\lambda)}{\partial\varphi_1}\right)\partial_w \tag{17}$$

These are the basis vectors of the tangent space in the coordinate basis system where (16) is essentially organized spatially in a matrix form and in (17) I explicitly take the matrix vectors and pull them out, (17) is the preferred way to think about it (16) is a sort of a common classical method of thinking about it. From all that I have now created one of the basis vectors of the tangent plane. I'm looking for all 3 but the other ones are pretty simple, the other one I just turn φ_1 to φ_2 , that's the value for $\vec{e}_{\varphi_2}$. λ goes from $(0,2\pi)$ and it shows up only in the last two of the transformation an if this were n dimensional it would still only show up in the last two. You can see how this angle element involves in that case so here I would have to now just turn φ_1 to λ , that's the value for $\vec{e}_{\lambda}$. We can create all 3 of these basis vectors by taking this prescription of derivatives so let's actually do that it's actually do that for the “Glome”.

$$\vec{e}_{\varphi_1}=\begin{pmatrix}-R\sin\varphi_1\\ R\cos\varphi_2\cos\varphi_1\\ R\sin\lambda\cos\varphi_1\sin\varphi_2\\ R\cos\lambda\cos\varphi_1\sin\varphi_2\end{pmatrix}=R\left[-\sin\varphi_1\partial_x+\cos\varphi_2\cos\varphi_1\partial_y+\sin\lambda\cos\varphi_1\sin\varphi_2\partial_z+\cos\lambda\cos\varphi_1\sin\varphi_2\partial_w\right] \tag{18}$$

$$\vec{e}_{\varphi_2}=\begin{pmatrix}0\\ -R\sin\varphi_2\sin\varphi_1\\ R\sin\lambda\sin\varphi_1\cos\varphi_2\\ R\cos\lambda\sin\varphi_1\cos\varphi_2\end{pmatrix}=R\left[-\sin\varphi_2\sin\varphi_1\partial_y+\sin\lambda\sin\varphi_1\cos\varphi_2\partial_z+\cos\lambda\sin\varphi_1\cos\varphi_2\partial_w\right] \tag{19}$$

$$\vec{e}_{\lambda}=\begin{pmatrix}0\\ 0\\ R\cos\lambda\sin\varphi_1\sin\varphi_2\\ -R\sin\lambda\sin\varphi_1\sin\varphi_2\end{pmatrix}=R\left[\cos\lambda\sin\varphi_1\sin\varphi_2\partial_z-\sin\lambda\sin\varphi_1\sin\varphi_2\partial_w\right] \tag{20}$$

For $\vec{e}_\lambda$ that's real easy, everything is zero except these last two and that will be true for any number of dimensions by the way because like I said for any number of dimensions λ only shows up in the last two I think I have that formula for n dimensions here:

$$\begin{cases} X^1=r\cos\varphi_1 \\ X^i=r\cos\varphi_i\prod_{k=1}^{i-1}\sin\varphi_k \text{ for } 1<i<n-1 \\ X^{n-1}=r\sin\lambda\prod_{k=1}^{n-2}\sin\varphi_k \\ X^n=r\cos\lambda\prod_{k=1}^{n-2}\sin\varphi_k \end{cases} \tag{21}$$

We've just calculated the 3 basis factors in the Euclidean space given any point in the “Glome” own coordinates. I should be a little bit more precise here and now's the time so we're dealing with these vector in 4D space the coordinate vectors in 4D space but I've got these these $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_\lambda$ objects as well and I want to make it real clear what the connection is, I don't know about clear but I want to make sure we struggle with what that connection is so if we live out in the 4D space so we're looking in at the “Glome” or likewise you might be able to think if we're living in 3D space and we're looking at the sphere equivalently and we use the (x,y,z) coordinate system, this is how we refer to all the tangent spaces at every point $(\partial_x,\partial_y,\partial_z,\partial_w)$ we use this coordinate system, this coordinate basis so any point in our 4D world we have access to these 4 basis vectors.

It turns out that when we stare at the sphere however we kind of realize we don't actually need all four of these because at any given point, well we need all four in some sense, but if I take one point on the sphere, one point given by $(\varphi_1,\varphi_2,\lambda)$ then I don't need all four, I can find a linear combination of these four or I can find three linear combinations of these four $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_\lambda$ and every vector in that tangent plane can be expressed with this basis so I can actually drop a dimension by taking the appropriate linear combination of $(\partial_x, \partial_y, \partial_z, \partial_w)$ and calling that the three basis vectors. Now that's a little bit annoying because I get outside of my coordinate basis system I'm no longer in the (x, y, z, w) coordinate basis system, on the other hand if I actually lived inside the sphere, if I lived inside the "Glome" my coordinate system is $\varphi_1, \varphi_2, \lambda$ and in that sense my fundamental coordinate system is going to be $\partial_{\varphi_1}, \partial_{\varphi_2} \partial_\lambda$ and it's only 3D it's not 4D and these guys $\partial_{\varphi_1}, \partial_{\varphi_2} \partial_\lambda$ are connected to these guys $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_\lambda$ through this relationship:

$$\partial_{\varphi_1} \equiv \vec{e}_{\varphi_1} \text{ , } \partial_{\varphi_2} \equiv \vec{e}_{\varphi_2} \text{ and } \partial_\lambda \equiv \vec{e}_\lambda \tag{22}$$

My coordinate basis vectors in "Glome" coordinate basis vectors $\partial_{\varphi_1}, \partial_{\varphi_2} \partial_\lambda$ are in fact what these guys $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_\lambda$ are. Any vector in "Glome" that I would write down some $\vec{X}$ is always going to be:

$$\vec{X} = X^{\varphi_1} \partial_{\varphi_1} + X^{\varphi_2} \partial_{\varphi_2} + X^\lambda \partial_\lambda \tag{23}$$

In General Relativity we would rename this φ_1 would be likely $\varphi_1 = X^{1'}$, $\varphi_2 = X^{2'}$ and $\lambda = X^{3'}$ and I would write $\vec{X} = X^{\mu'} \partial_{\mu'}$ like that and that would be in "Glome" coordinates and the out of "Glome" coordinates would be $\vec{X} = X^\mu \partial_\mu$. Those are two different coordinate systems, in this case, in order to keep the 4D I probably have to keep r as a coordinate otherwise $\vec{X} = X^{\mu'} \partial_{\mu'}$ would only have three, $\vec{X} = X^\mu \partial_\mu$ would have four but the point is that naming the coordinates independently is a convenience for this kind of analysis. The point is that the coordinate basis in "Glome" is related to the coordinate basis out of "Glome" by this relationship (14) which you could invert by the way. You can invert it but if you inverted it you could still only capture these sub spaces that are on the "Glome".

We are going to work on this induced metric finally so the point is now and I'll use the sphere as our example if I have this if I'm dealing with vectors in the tangent plane of a sphere and I call one of those vectors $\vec{X}$ and one of those vectors $\vec{Y}$ then I have in principle two ways of calculating the inner product of $\vec{X}$ with $\vec{Y}$. I can use the metric of the underlying Euclidean space G_E or I can calculate the inner product of $\vec{X}$ with $\vec{Y}$ using the metric of the underlying sphere G_S , we don't know the metric of the underlying sphere but we do know the metric of the underlying Euclidean space or in our 4D case this would be the "Glome" G_S and this would be $\mathbb{R}^4$. If I express $\vec{X}$, $\vec{Y}$ in terms of the coordinate basis of $\mathbb{R}^4$, $(\partial_x, \partial_y, \partial_z, \partial_w)$, I can use this metric G_E to calculate the inner product of $\vec{X}$ and $\vec{Y}$.

Likewise if I expressed $\vec{X}$ and $\vec{Y}$ in terms of the coordinate basis of the "Glome" itself $\partial_{\varphi_1}, \partial_{\varphi_2} \partial_\lambda$ then I would have to use this metric G_G which I don't know but the good news is that regardless of which metric you use the presumption is the answer is going to be the same, $\vec{X}$ and $\vec{Y}$ has an inner product and that inner product should be independent of whether or not I use the coordinate system attached to the "Glome" or the coordinate system attached to the embedding space. I'm going to use that fact to calculate this metric. What I would normally do is I would say I've got:

$$\left(X^\mu \partial_\mu, Y^\nu \partial_\nu\right)_E = \left(X^x \partial_x + X^y \partial_y + X^z \partial_z + X^w \partial_w, Y^x \partial_x + Y^y \partial_y + Y^z \partial_z + Y^w \partial_w\right) \tag{24}$$

Then by linearity, what do we end up with:

$$\left(X^{\mu}\partial_{\mu},Y^{\nu}\partial_{\nu}\right)_E=X^xY^x\left(\partial_x\partial_x\right)+X^yY^y\left(\partial_y\partial_y\right)+X^zY^z\left(\partial_z\partial_z\right)+X^wY^w\left(\partial_w\partial_w\right) \tag{25}$$

There are no cross terms because the cross terms are zero in because we have:

$$\text{Euclidean}\rightarrow\ \left(\partial_{\mu},\partial_{\nu}\right)=\delta_{\mu\nu} \tag{26}$$

That inner product has to equal the inner product $\left(X^{\mu'}\partial_{\mu'},Y^{\nu'}\partial_{\nu'}\right)_G$, Now we actually have some trouble because this is going to end up being sort of matching μ' to $\varphi_1,\varphi_2,\lambda$

$$\left(X^{\mu'}\partial_{\mu'},Y^{\nu'}\partial_{\nu'}\right)_G=\left(X^{\varphi_1}\partial_{\varphi_1}+X^{\varphi_2}\partial_{\varphi_2}+X^{\lambda}\partial_{\lambda},Y^{\varphi_1}\partial_{\varphi_1}+Y^{\varphi_2}\partial_{\varphi_2}+Y^{\lambda}\partial_{\lambda}\right) \tag{27}$$

This is going to be:

$$\left(X^{\mu'}\partial_{\mu'},Y^{\nu'}\partial_{\nu'}\right)_G=X^{\varphi_1}Y^{\varphi_1}\left(\partial_{\varphi_1},\partial_{\varphi_1}\right)_G+X^{\varphi_1}Y^{\varphi_2}\left(\partial_{\varphi_1},\partial_{\varphi_2}\right)_G+\cdots \tag{28}$$

We're not sure about the cross terms. This is in the “Glome” coordinates. Ultimately all of this stuff, all of these guys have to be equal and in the coordinate basis. We want to execute this calculation so we can learn what these these guys are $\left(\partial_{\varphi_1},\partial_{\varphi_1}\right)_G$, we're after these guys because these guys that would be $g_{\varphi_1\varphi_1}$ because the definition of $g_{\alpha\beta}$ is:

$$g_{\alpha\beta}\overset{\text{def}}{=}\left(\partial_{\alpha},\partial_{\beta}\right) \tag{29}$$

That's the definition of what these metric coefficients are so we're looking to capture and figure out what those guys are, but if that's all we're doing if we're just trying to calculate what these guys are $\left(\partial_{\varphi_1},\partial_{\varphi_1}\right)_G$ to get these metric coefficients, well let's calculate them I mean we know the relationship between $\varphi_1,\varphi_2,\lambda$ and the Euclidean basis of the underlying space (14). To demonstrate let's do:

$$g_{\varphi_1\varphi_1}=\left(\vec{e}_{\varphi_1},\vec{e}_{\varphi_1}\right)_E=\left(\partial_{\varphi_1},\partial_{\varphi_1}\right)_G \tag{30}$$

When I wrote down these $\vec{e}_{\varphi_1}$, $\vec{e}_{\varphi_2}$ and $\vec{e}_{\lambda}$, I wrote them in terms of the Euclidean basis (18), (19) and (20). What I'm going to do now is I'm going to make the substitution, I'm gonna do calculate this in Euclidean space and I'll get a number and that number will be the coefficient. These are two Euclidean vectors and this is pretty easy, now I take the inner product of $\vec{e}_{\varphi_1}$ with $\vec{e}_{\varphi_1}$ using linearity and knowing that all the cross terms are zero because of (26), the only ones that survive are the diagonal ones so the answer is going to be:

$$g_{\varphi_1\varphi_1}=R^2 \tag{31}$$

This is the surface of the sphere, is gonna have a much more complicated way of figuring out how to find the distance between points it's not going to be nearly as simple as the Euclidean metric it's gonna be something else and clearly you can see it's gonna be some complex function of $\varphi_1,\varphi_2,\lambda$.

It is actually kind of nice and simple for that particular first component so we can do this for the 2nd component, we can look for:

$$\mathrm{g}_{\varphi_2\varphi_2}=\left(\vec{e}_{\varphi_2},\vec{e}_{\varphi_2}\right)_E=\left(\partial_{\varphi_2},\partial_{\varphi_2}\right)_G=R^2\sin^2\varphi_1 \tag{32}$$

I guess we're left with the calculation of:

$$\mathrm{g}_{\lambda\lambda}=\left(\vec{e}_{\lambda},\vec{e}_{\lambda}\right)_E=\left(\partial_{\lambda},\partial_{\lambda}\right)_G=R^2\sin^2\varphi_1\sin^2\varphi_2 \tag{33}$$

Now I'll leave you guys to prove for yourself that all the cross terms are zero so all the non diagonal components are zero because there's symmetry. We have found all of the matrix elements and therefore

$$g_{\alpha\beta}=\begin{pmatrix} R^2 & 0 & 0 \\ 0 & R^2\sin\left(\varphi_1\right)^2 & 0 \\ 0 & 0 & R^2\sin\left(\varphi_1\right)^2\sin\left(\varphi_2\right)^2 \end{pmatrix} \tag{34}$$

You could also write it down in metric form which would be:

$$ds^2=R^2\left(\sin\left(\varphi_1\right)^2\mathrm{d}\left(\varphi_2\right)^2+\mathrm{d}\left(\varphi_1\right)^2+\sin\left(\varphi_1\right)^2\sin\left(\varphi_2\right)^2\mathrm{d}\left(\lambda\right)^2\right) \tag{35}$$

Which could also be written in the form:

$$ds^2=R^2\left[d\varphi_1^2+\sin^2\varphi_1\left(d\varphi_2^2+\sin^2\varphi_2 d\lambda^2\right)\right] \tag{36}$$

This guy here (36) is sort of the normal form of this metric and it looks very much like the standard spherical form of a metric, the spherical form of the metric was:

$$ds^2=R^2\left(d\theta^2+\sin^2\theta\, d\varphi^2\right) \tag{37}$$

This is sort of the same λ takes the place of φ and φ_1 takes the place of θ , you have a $\sin^2\theta$, etc. What you're missing is φ_2 but you don't have a 2nd $(0,\pi)$ angle on the regular sphere. It just gets a little more complex but you can imagine that (36) is like the most elementary version of something that looks like (37) which is a little more elementary than the next dimension which will have an additional $(0,\pi)$ type angle. This is the metric on the “Glome” (36) and the reason I went through this trouble to get the metric on the “Glome” is if you look at the spatial part of the conform-compactified coordinates metric, (2.1.4) in the catalog, that's exactly what you're dealing with. We have two angles $\xi\in(0,\pi)$ and $\theta\in(0,\pi)$ and then you have one angle $\varphi\in[0,2\pi)$ and this one $\psi\in[-\pi,\pi]$ is the time variable, just ignore that for now because we're just looking at the purely spatial Euclidean part and that's the “Glome”. Ultimately what we learned is that Minkowski space, when we did that whole analysis of Minkowski space the spatial part of Minkowski space in compactified coordinates is conformal to the

Lorentzian “Glome” and that's kind of why I chose this example so I want to go through this catalog with some thoroughness, I wanted to understand what this is, this is the 3-sphere or the “Glome” and $-d\psi^2$ is a time coordinate and Minkowski space is conformal to that in the sense that the metric of Minkowski space when you compactify the coordinates the way we did it's related to this “Glome” metric through the simple multiplication by a function (2.1.17) and we haven't talked about conformal coordinates yet but (2.1.15) is not Minkowski space, (2.1.17) is Minkowski space, the relationship between the two is something called conformal relationship and it's interesting that Minkowski space has this conformal relationship to the 4D “Glome”.

We showed how you can take an embedded surface and discover the metric of the embedded surface by utilizing the metric of the embedding space and in this case it was Euclidean space, the point being that the Euclidean space is simple, we understand its metric very well and now we've gone through this whole massive process and we learned that the embedding space, its metric it's going to be somewhat complex but you could calculate the length of curves as long as it's in the subspace, in the sphere or in the “Glome” using the metric of the coordinates associated with just the sphere or the “Glome” and the point is that in General Relativity we don't do this, ironically, in General Relativity we have 4D space but we don't think of 4D space as being embedded in some sort of 5D space. We're going to use only the coordinates within the space itself, it's as though we live on the “Glome” and our metrics are going to be kind of complicated, they're not complicated because we're in a space sitting in a higher dimensional space, they're complicated because the metric is a distorted wacky version of flat space-time because of the Einstein equation.

Now there's this temptation to think well maybe there's a simpler metric in a higher dimensional space, that could make everything easy but we don't do that and we don't even try to do that and it's not clear that there is because we're not a smooth nice “Glome” anyway, we're some wacky shaped space-time but we're stuck with these hard metrics in our space-time but we're looking at it from inside our own space-time we're not referring it as though we're embedded in some higher dimensional space. It's like we're stuck on the “Glome” and we have to live with this metric here, this complicated metric to figure out how far apart things are.

Anyway now we've gotten a little facility with metrics and next time we put together we're gonna start attacking this problem here this question (1.3.3), we're gonna start discovering how the connection is related to the metric and that's actually quite a long journey.

Lesson 9: Parallelism and the Covariant Derivative

We are going to move on into the Catalog of Space-times. I've decided this is gonna basically be the sort of baseline reference book that we're gonna work with and the idea is to sort of attack it so that you understand it and if we really peel apart all of the pieces of the catalog of space-time you'll understand this is our approach to General Relativity since we're not going to be working out of a textbook and we need something to guide us that we can all work off of this is what that's going to be. Like I pointed out you can get this thing here at this URL indicated down here.

This first picture, this very first picture at the cover of the catalog of space-time that looks a lot like this drawing that I always have been drawing during all of the “What is a tensor” lectures, this thing about the different coordinates where I put these two spare coordinates up here, they're kind of doing the same thing this M represents the idea that space-time is modeled by a manifold which we discussed at length in the prerequisite material for this course and as a manifold we know by definition you can put down coordinates on the manifold because that's what a manifold is, it allows you to put down coordinates and here are two examples they use X^1 and X^2 , I usually use X^0 and X^1 when I do my little sketches and interestingly what they're doing here is they're taking each point in the manifold and they're kind of trying to lay down these vectors and this is the common technique that's used in almost every book out there is they want to say that at this point you have a vector space and what they've done in this little sketch where you see the coordinate basis, is what they're using, and the coordinate basis is, as we know, is a differential operator where it would take where it would operate on functions on the space-time so if $X\in\mathsf{M}$, some function of X , you can operate on that thing with a differential operator, in this case ∂_{x^1} which actually that's is kind of weird notation it's usually just ∂_1 , the x is understood, but they've got the superscripted subscript problem but it looks fine and what they're showing here is the basis vector of that process where it's sort of aligned tangentially with the line of constant coordinate value in this case coordinate value $x^2=0$.

Our tendency wouldn't be to do this, the way we understand all this is slightly differently. When we do this, we wouldn't actually draw arrows on the space-time because we're kind committed to this idea that at every point there's a separate tangent space so I would take say this point and I would say can pull it off and I would generate a whole new little object that would be the tangent space of, I guess in this case it would be $x^1=0$, $x^2=2$ so the point is $T_{(0,2)}(\mathsf{M})$, a tangent space on the manifold and inside this tangent space I would put these basis vectors, well let me just do it my way ∂_1 and ∂_2 and I would do this in every spot so this spot here would get its own tangent space in that case it's $T_{(1,0)}(\mathsf{M})$ and inside this tangent space the basis vector will be ∂_1 and ∂_2 that exists at this point and these two things are completely separate. Our conversation today is going to revolve around that separateness because we are going to attack in the very first part of the catalog section **1.3 Basic objects of the metric**, we are going to start attacking this relation (**1.3.1**) and this relation (**1.3.3**) basically we're going to attack all of these, all four (**1.3.1**), (**1.3.2**), (**1.3.3**), (**1.3.4**) of these we're gonna begin that attack today, by attacking I mean we're going to start to understand what these things are.

The idea is we now need to start talking about the connection on the manifold so the connection that we have studied is actually this object here drawing this object (**1.3.3**), this thing is gonna be related to (**1.3.1**) it but it's different because its index structure is different. We're going to begin our conversation about what's on the left-hand side and on the right-hand side we have a bunch of things involving the metric and now the metric, we know that the metric say in this case $g_{\rho\nu}(x)\, dx^{\rho}\otimes dx^{\nu}$. We know that the coefficient is a function of space-time, $g_{\rho\nu}(x)$ this is a $(0,2)$ rank tensor.

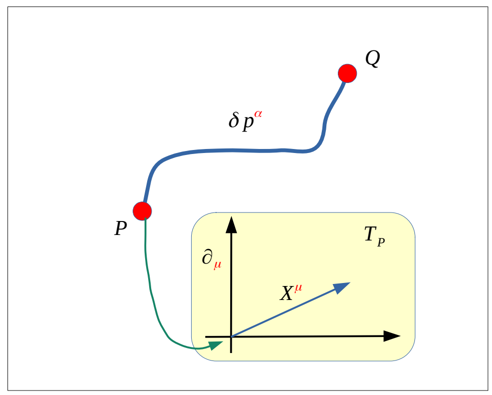
Therefore $g_{\rho\nu,\lambda}$ is simply $g_{\rho\nu}$ that thing (I'm going to suppress the argument of space-time, you always understand these things to be fields unless specified) is the partial derivative of that with respect to λ which means where λ could take any value like ρ and ν but this is a simple partial derivative that's all that says it's a simple partial derivative we know it's a simple partial derivative because this symbol here is a comma, if it was a semicolon that's the standard notation for the Covariant derivative which is something we discussed in our prerequisite material but we're gonna go through again today in great detail, but this is simply a partial derivative and it's not a tensor because we know simple partial derivatives of tensor components do not yield other tensor components. They don't transform correctly but that makes sense because we already know that this thing on the left-hand side is not a tensor, famously, so there's no reason to think the thing on the right-hand side is, in fact it's an equality so they can't both be tensors but it makes perfect sense so these three things here $g_{\rho\nu,\lambda}$, $g_{\rho\lambda,\nu}$, $g_{\nu\lambda,\rho}$ these are all partial derivatives, $g^{\mu\rho}$ is the metric tensor components and this combination of partial derivatives of the metric multiplied by the metric tensor component like (**1.3.3**) where the Einstein summation convention is operating on ρ and μ,ν,λ are the indices that are left over. Ultimately I need to explain to you why this is true, I need to explain to you why the metric when combined in this way gives you a connection on the manifold and that's our goal for the next few lessons.

Let's begin by reviewing what the connection actually is, and there's this geometric picture that in the elementary methods of doing this stuff is used in so many textbooks that it's totally worth knowing and understanding and connecting to the abstract view so we're going to start with this basic geometric view of the connection and it's the typical picture in the textbook, would look like this a point P and a point Q . This whole thing is a manifold so there's a coordinate system, when I write down P so that's an important thing to understand when I write down P as a point what I mean is, it is a point in the manifold but it also has coordinates and those coordinates are $P=\left(X^0(P),X^1(P),X^2(P),X^3(P)\right)$.

From your understanding of manifolds, you understand that these are really charts these are charts that take P from the underlying manifold and give you a real number and the overall chart which would be $X(P):\mathsf{M}\to\mathbb{R}^4$ is a map from the manifold to $\mathbb{R}^4$, that's just what manifolds are so when I write down P and when I write down Q like this you need to think it's an object with coordinates and those coordinates although I'm writing P down you it's totally fine to think of it in terms of coordinates because it's a manifold that's the whole point of having a manifold is it allows you to freely think of P and Q as objects with real value coordinates in $\mathbb{R}^4$, that's because our manifolds in General Relativity are 4D manifolds, that's part of the sort of axiomatic nature of General Relativity is we have a 4D manifold and the reason we have a 4D manifold is we have the Equivalence principle which tells us that on every 4D manifold of General Relativity we can find a coordinate system where the metric on the manifold is Lorentzian, so it kind of goes back to the axioms of Special Relativity which invented the idea of a 4D flat manifold as the place where you can study Electromagnetism and dynamics and all kinds of stuff so that's where the 4D come from.

We have these two points P and Q , now I'm going back to my explanation of what is the connection $\Gamma^{\alpha}_{\beta\gamma}$ so this is a review of what we did in the “What is a tensor” series I forget which precise lecture discussed it, It's worth going through again, you can't go through this stuff too many times I have no fear of redundancy as I think anybody who's paving the attentions lectures probably knows by now. You've identified two points in space-time, let's put it that way, and you have a vector inside the tangent space at P now the way textbooks always do this is they just draw it like this they just say here's our vector, it's at the tangent space of P and to signify this they put the tail of the vector at P but everybody knows that this vector does not exist inside the space of the manifold, instead it exists inside the tangent space at P .

The way we would do it, the way I would be inclined to do it, is I would be inclined to say: here's the tangent space at P at some other space, it's called $T_P(\mathsf{M})$, that's the definition of this vector space and it has its basis vectors γ_μ and associate with it is the Cotangent space $T^*_P(\mathsf{M})$ and then associated with that or all of the different possible tensor product spaces that might exist at P built of vectors and co vectors from the tangent space and cotangent space at P , that's how I would do it if I were writing a textbook I would try to make the picture look like that because I really feel like this (the first page of the catalog) gets a little confusing. Now it turns out that the Equivalence principle does eventually allow us to write vectors this way and that's why this works because remember the Equivalence principle asserts that at P there is a flat space-time and at a flat space-time parallelism is very clear and in fact we're gonna create a construction little later called Schild's ladder which shows us how to construct geodesics using this notion of the Equivalence principle. I don't know when we'll get to it, I don't know when we'll get to it but we will soon. If I were to do this I would put that vector inside this vector space and I guess maybe I would draw some axes to reflect the orthogonal basis ∂_0 , ∂_1 , ∂_2 because there is some coordinate system on this manifold and the coordinates are $X^0, \ldots, X^3$ so this vector has components inside the tangent space. Now you can't draw those components here I can't do this at the point because then I would have to pretend like those basis vectors exist on the manifold and that gets awkward.



With that in mind, I want to ask the question: say I went from P to Q along some path and we make the path, there's no reason not to make the path kind of straight and that path I'm gonna call δp^α and that path takes me from P to Q and it's going to be a small amount of distance which is a bit tricky to talk about because if I say it's a small displacement, if I use that language small displacement I have two problems: I have the problem of small and I have the problem of displacement.

The problem with small is what do I mean by small when I don't have a metric, I haven't said that this is a metric space necessarily, now in General Relativity everything is a metric space so we all we do have a metric but what I'm going to talk about now is just the connection and the connection can exist separately with or without a metric, in other words if I have a metric I am eligible to define a connection using these formulas (**1.3.3**) which we will eventually derive so if I have a metric then I know that I can create a certain kind of connection that has a certain set of properties so that's good but if I don't have a metric I can still define and create a connection I just have a different connection it's not the metric connection it's some other Affine connection so that's what I'm talking about here, I'm talking about a situation where I just have a connection because we want to understand the connection in and of itself so when I go back to this I can say I understand why I need a connection $\Gamma^\mu_{\nu\lambda}$ and why connections are important and this is really great given that my space-time has a metric then the good news is that I will automatically get a connection from that metric because the physics gives us the metric therefore the physics gives us this all-important object called the connection.

Going back to this, the point of that is that the notion of small is difficult because I don't necessarily have a metric to decide what is small and large and to compare things with so we kind of have to address that point and then displacement. Displacement sounds a lot like a vector quantity but this thing is just literally a change in coordinate values so when I write down δp^α :

$$\delta p^{\alpha}=(\delta p^0,\delta p^1,\delta p^2,\delta p^3) \tag{38}$$

I'm thinking about the change in all the coordinates, I'm thinking about the changes in 4 coordinates and that takes me from (P^0,P^1,P^2,P^3) those are real numbers that define this point P on the manifold the alternative is $(X^0(P),X^1(P),X^2(P),X^3(P))$ so I'm just gonna call it (P^0,P^1,P^2,P^3) for simplicity say likewise this guy will be $(X^0(Q),X^1(Q),X^2(Q),X^3(Q))$ or (Q^0,Q^1,Q^2,Q^3) , maybe I can get away with that so the point is that:

$$Q^0=P^0+\delta p^0 \quad , \quad Q^1=P^1+\delta p^1 \quad , \quad Q^2=P^2+\delta p^2 \quad , \quad Q^3=P^3+\delta p^3 \tag{39}$$

That's what this displacement means. Now the question of small, that is a little bit more subtle, it turns out not to be a problem because we're on a differentiable manifold we have access to calculus and ultimately the idea is we're gonna drive δp^{α} to zero in a limit and as we drive it to zero in a limit it will automatically become small because there's no definition of small that isn't satisfied if you drive this to zero in the limit so the idea is that we're using limits to zero as being small so all you need is to have solid differentiable manifold structure which includes all the topology of the manifold and that's what what gives us our notion of small ultimate.

What's the question we're trying to answer? I've got a vector which I called X^{μ} . It's really a [Fiber bundle] where this is the [Tangent bundle] of the manifold where the tangent bundle being the collection of the manifold and all of the tangent spaces attached to the base space, you may remember that from the prerequisite material on “What is a manifold” but if you just think about this is the tangent space T_P associated with this point P and the vector in that tangent space then I think this is a picture that's what this is a vector inside this tangent space which is associated with that point P and like I said most textbooks just like to say: forget all this, I'm not going to draw a whole separate space I'll just attach the vector to the point and try to illustrate it that way.

My question though is if I were to take that vector and “move” it, well I guess I'll put it this way, let's do it our method: at Q there is a tangent space and it has basis vectors $\partial_0,\ldots,\partial_3$ so there's some vector in this tangent space whose components I don't know but that vector is going to be understood to be parallel to X^{μ} so I need to find a vector in Q that's parallel to X^{μ} and the obvious problem is, I'm trying to establish a relationship between the members of this tangent space T_P and the members of this tangent space T_Q . Remember they're totally different vector spaces so there is absolutely no connection between them until you make one, until you define one, until you invent one, until you create a connection and the reason I keep saying all these words, define, invent, create is because there's no natural connection there's no canonical connection between any tangent spaces in a manifold, you have to impose one arbitrarily. One choice that you have, if there's a metric involved, is you always will have this choice (**1.3.3**) which will study in detail as we go on I keep promising that and we will get there but you always have that choice so if you do have a metric you are guaranteed that you have (**1.3.3**) as a potential connection, you don't have to use it though, you can put any connection in the manifold it just so happens this is a very logical choice but it's not the only choice, I could just say no I've got this metric and I'm using some other connection but there's a very good reason in Physics to use this connection (**1.3.3**). Now we want to understand what does how do you use this connection to create this notion of a vector in one tangent space being parallel to a vector in another tangent space and it's not that hard concept.

The tangent space at P and the tangent space at Q had their basis vectors in the coordinate basis. Since we have a coordinate system, we're allowed to have a coordinate basis, remember the distinction between the coordinate basis in P tangent space T_P and this one in Q, T_Q is simply that when it acts as a differential operator, it's evaluated it's evaluated at the coordinates of P and this one's evaluated at the coordinates of Q so they actually are different. In T_Q I'm going to create a vector that is parallel to X^{μ} and I'll call that $X^{\parallel\mu}(P,Q)$ meaning it's the vector in T_Q that's parallel to the vector $X \in T_P$ and the components of the thing are μ and then I have the displacement δp^{α} like that.

When I assert is that the components of X^{μ} are kind of close to the components of $X^{\parallel\mu}(P,Q)$ meaning that the zero component of X and the zero component of $X^{\parallel\mu}(P,Q)$, the difference between those two is going to be small and we could call that:

$$X^0 - X^{\parallel 0}(P,Q) = \bar{\delta} X^0 \tag{40}$$

This $\bar{\delta} X^0$ represents the small change in the component value between X^0 and $X^{\parallel 0}(P,Q)$ and small in the sense that as $\delta p^{\alpha} \rightarrow 0$ then $\bar{\delta} X^0 \rightarrow 0$ and this is telling us that however we define parallel it has to be a relatively smooth concept mathematically speaking and by mathematically speaking we mean that in the limit that the coordinate difference between P and Q that's this thing $\bar{\delta} X^0$ goes to zero the difference in the components between this parallel vector and the original vector have to get smaller and smaller and indeed if Q was actually at P this thing would be zero because a vector would be presumed parallel to itself in its own vector space so in that sense we can define small in a very strict and calculus based sense, everything needs to be smooth and continuous for example when we do calculus we don't really define metrics per se but we made sure that everything is continuous and smooth and that's what we're talking about here.

It all comes down to this thing $\bar{\delta} X^0$, can we figure out what is the change in components given a small value of δp and that is how we define the connection. We begin by saying: whatever this thing is, whenever this small value of the change in the component is, it's going to be proportional, let's say it's proportional to δp^{μ}, it's proportionate to the displacement and it should also be proportional to the magnitude of the component itself so if the displacement is zero then $\bar{\delta} X^{\mu}$ should be zero and if the vector component itself is zero then $\bar{\delta} X^{\mu}$ should be zero. We want those two things to be proportional $\bar{\delta} X^{\mu} \propto \delta p^{\nu}$ and $\bar{\delta} X^{\mu} \propto X^{\rho}$. We're going to call it linear just to keep it kind of simple, there's no way to over presume, over complicated kinds of relationships so with that we're gonna say it's proportional to the displacement, it's proportional to whatever the displacement is because the 0^{th} component of this change is proportional to displacement but it can be proportional to any direction of displacement, it can be proportional to the 0^{th}, 1^{st}, 2^{nd} or 3^{rd} proportion of displacement. I need to put it different index up there so it's gonna be proportional to all of those, whatever the constant of proportionality is, it's going to be proportional to all of those things, all of the different little displacements, the displacements in each of the different coordinate directions. Likewise with the component of X, it's going to be proportional to the component in different directions so the way to think about this is that the difference in the say $\bar{\delta} X^0$ component it's proportional one way or another to:

$$\begin{gathered}\bar{\delta} X^0 \propto \delta p^0 \;,\; \bar{\delta} X^0 \propto \delta p^1 \;,\; \bar{\delta} X^0 \propto \delta p^2 \;,\; \bar{\delta} X^0 \propto \delta p^3 \\ \bar{\delta} X^0 \propto X^0 \;,\; \bar{\delta} X^0 \propto X^1 \;,\; \bar{\delta} X^0 \propto X^2 \;,\; \bar{\delta} X^0 \propto X^3\end{gathered} \tag{41}$$

It's proportional to all of those things. Let's write down that:

$$\bar{\delta} X^{\mu} = -\Gamma^{\mu}_{\rho\nu} X^{\rho} \delta p^{\nu} \tag{42}$$

Where $\Gamma^{\mu}_{\rho\nu}$ is some constant of proportionality that sums over ρ and ν . There's a minus sign here for convention, that's a convention, this connection coefficient, the connection represents difference in components between the vector X^{μ} which is at P and the vector which we called $X^{\parallel\mu}(P,Q)$ so the components of the parallel vector infinitesimally displaced in the space-time from the vector in the tangent space of P , those components change by this much $\bar{\delta} X^{\mu}$ and in order to establish what that much is, we need this connection coefficient $\Gamma^{\mu}_{\rho\nu}$ and we need to know the magnitude of the original vector component and we need to know the little displacement and you'll notice if you displace in different paths, you're gonna get a different answer. Parallelism is path-dependent in this circumstance.

If P is here and Q is here I can imagine a path for P to Q made up of a lot of little infinitesimal displacements and if I take this path versus this path I'm going to end up with a different parallel vector that's a very famous principal here. Now if you take the simplest case that a flat space-time we know that from elementary physics if I have a vector I'm free to transport that vector over here but I'm still doing the same thing because T_P is still a tangent space and the tangent space has a vector and that vector is X^{μ} and T_Q is a tangent space and presumably that tangent space has some parallel vector which we were calling $X^{\parallel\mu}(P,Q)$ but our rule was that these things are the same, there's no change in the components, the components are always the same which basically means $\bar{\delta} X^{\mu}$ is always zero which means that $\Gamma^{\mu}_{\rho\nu}=0$, the connection is zero in flat space-time and that's why we can take a vector in flat space-time and just move it from here and slide it to there and we can let it sit in the space-time plane also very comfortably without having to do this picture of a separate tangent space although we could, all we need to know is that the components don't change.

If we change to Spherical coordinates then the components would change, you would still slide the vector but now you have to recalculate what the components are and indeed the metric and the connection in Spherical components, as we showed, I think in a previous lecture back in the "What is a tensor" series, they are not all zero so when you do a coordinate transformation you can take things that are all zero, you can take a connection that's all zero in one set of coordinates and you can produce, through a coordinate transformation, a connection that doesn't have all zero values in another coordinate system.

Notice for a tensor you couldn't do that, a tensor if this was a true tensor $\boldsymbol{T}^{\mu}{}_{\rho\nu}$, first of all you wouldn't stack these symbols on top of each other, you would have to separate them because otherwise you can't keep track of the order but if you had a tensor and all the components were zero in one system well the coordinate transformation equation involves:

$$\boldsymbol{T}^{\mu'}{}_{\rho'\nu'} = \frac{\partial x^{\rho}}{\partial x^{\rho'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial x^{\mu'}}{\partial x^{\mu}} \boldsymbol{T}^{\mu}{}_{\rho\nu} \tag{43}$$

It's all products and if $\boldsymbol{T}^{\mu}{}_{\rho\nu}$ is always zero doesn't matter how many products you have you're always going to get zero so a real tensor that zero on one coordinate system will be zero in every coordinate system but $\Gamma^{\mu}_{\rho\nu}$ does not transform with this rule, this rule here is only for tensors, this is standard coordinate transformation for tensors. $\Gamma^{\mu}_{\rho\nu}$ has a totally different rule which by the way it's the same it has part that's identical for tensors but then it has another part and that other part is what allows you to get nonzero terms from zero terms. I think we reviewed the coordinate transformation rules for connections in the previous lecture but the point is, this is what the connection is doing, the connection is defining for us parallelism and you can move from an infinitesimal displacement to a finite displacements by integration and through that you can find a parallel of any vector from a tangent space at one point to the tangent space of have another point.

In order to do it you have to define the connection on the manifold so we'll talk about that process for finding parallel vectors anywhere in the manifold given a path and a parallel vector will do that but I want to run through the rest of the concepts that all of this is tied to and this diagram that we see in so many textbooks where we start with a vector X and we talk about the components of X or X^{μ} and at the point Q . I'm using the idea that normal parallelism, we would have considered to be some vector like that parallel on the plane but I'm gonna say well no, in the world of manifolds with arbitrary geometries this parallel vector will have a different set of components and those components will be $X^{\mu}+\bar{\delta} X^{\mu}$ and this is the vector at Q that is truly parallel to this vector X^{μ} at P so the idea of [Parallel transport] is that we're taking a vector at P and we're moving it along some curve and at every point, as it moves, what we've done is we've just picked out the curve at this Q , let's call this point R and maybe this point S , we pick the vector at R that is parallel to this initial vector X^{μ} . Now I can only write this form for an infinitesimal displacement but we start talking about finite displacements then I start writing things like $X^{\parallel\mu}(P,R)$ that's is the symbol I was using, and then this one here would be $X^{\parallel\mu}(P,S)$ so there's some vector here. The process of Parallel transport is the process of this calculation of starting with this vector X^{μ} and then calculating this vectors components $X^{\parallel\mu}(P,Q)$ and this vectors components $X^{\parallel\mu}(P,R)$ and this vectors components $X^{\parallel\mu}(P,S)$ and then placing it in each of these tangent spaces.

It begins with a vector at some point in the space-time in this case is this vector X^{μ} and a connection, and the connection is a function of the space-time and if I know X^{μ} and I know this connection $\Gamma^{\mu}_{\rho\nu}$ I can create this process of Parallel transport and I'm warming up the concept now but we will actually execute this and do an example probably in the next lesson so with the concept of Parallel transport starting from a single vector I can define a vector anywhere along some path in space-time and I can do this for any path between any two points and I do understand that the vector I get at this point will depend on the path I take. $X^{\parallel\mu}(P,R)$ is not enough and then I need some indication of what the path is I would need to add $X^{\parallel\mu}(P,R,\gamma(\tau))$ Where γ is the path and τ is a parameter so $\gamma(\tau)$ is a path, it is a curve in space-time. I need to define that too so the name of this parallel vector has got to include the starting point the name of the actual vector that lives at that starting point the ending point and the path you took and then I can give it a component so $X^{\parallel\mu}(P,R,\gamma(\tau))$ is just a name of something.

From this we can now develop the concept of a Covariant derivative, though that's what's important about this, we can develop the idea of the Covariant derivative and this is where the notation in different books really goes haywire. The concept of the Covariant derivative exploits this notion of the path in the following way and this is again we covered all this in the previous lecture series “What is a tensor” but I'm going through it again because you can't you really can't overdo this stuff.

We start with a point P and we go to a point Q and this is the vector X that's attached to the tangent space of the point P and this would be the components $X^{\mu}=(X^0,X^1,X^2,X^3)$. Now here's the thing about X , X we're gonna take to be a Vector field so the vector X is a Vector field so there is a vector X at every point and this thing is just $X(P)$ but there's not only an $X(P)$, there's also $X(Q)$ which most books will just attach it to Q and draw it like this X at the point Q .

X is a Vector field so a Vector field takes a point in the manifold and returns a vector and these Vector fields are used in many different ways for different physical problems, they model Electromagnetic fields they model all kinds of physical processes flows of fluids they could they could model a lot of stuff so let's not worry about exactly what the Vector field is or what it represents, it's mathematically a Vector field and it's a tensor, a Vector field is a tensor so $X(x)$, X is the name of the Vector field, while x is a point in space-time so that's unfortunate but I'm going to stick with it, that is going to be $X^{\mu}(x)\partial_{\mu}$, it's a tensor, it's a $(0,1)$ Tensor field because the component is a function of space-time and it's given in the coordinate basis so that's what I mean by these symbols like $X(P)$ so when you see $X(Q)$ here you can think of $X^{\mu}(Q)\partial_{\mu}$ like that. Likewise here $X^{\mu}(P)\partial_{\mu}$ that's what those things actually are but again just to keep in mind because I'm trying to prepare you for examining books you rarely see them carry this around in many General Relativity textbooks and they generally do not carry the arguments around either and so that's all you see X^{μ} at P and then I guess you would have to have some indication of what it is at Q . I guess you would have to keep the P maybe they would put a P as the subscript and Q as a subscript or something, not to be a space-time index but to be some indication of what tangent space it is but since we do seem to need that I'm gonna just I'll leave the argument in here. Actually what they would write in this case, you see in textbooks as well, what's the distance between P and Q , as I said before they give it as δp^{α} that's the displacement.

Then we're going to ask the question is what kind of derivative can we take of this Vector field and we want the derivative of this Vector field to be some kind of tensor object because tensor objects are kind of what physics is all about, it's got to be something that doesn't depend on a coordinate system, if you have a definition of the derivative that depends on the coordinate system you're using that doesn't help much but you still have the basic question, well you've got a vector in one tangent space here, this is the same problem as we always have and then you have this vector in another tangent space you can't compare them but what you can do is you can compare this vector here $X^{\mu}(Q)$ which is a small displacement away from P , not to this vector here $X^{\mu}(P)$ but you can compare it to this vector Parallel transported to Q because there's a vector in the tangent space of Q , we've just been talking about it and that vector has components at Q of $X^{\mu}(P)+\bar{\delta}X^{\mu}(P)$ and that, this object here, we gave sort of the full kind of weird name $X^{\parallel\mu}(P,Q)$ (I'll just leave out this issue of the path for now and actually in this infinitesimal distance the path will become irrelevant anyway).

This parallel vector here, that is a vector in the tangent space at Q . Notice that $X^{\mu}(P)$ is a component of a vector in the tangent space at P and $\bar{\delta}X^{\mu}(P)$ is a change in the component of a vector in the tangent space of P but when we add them together we're putting them on a vector in the tangent space at Q so we've done something weird, we've associated a vector in Q for the vector in P but I guess I've said that a million ways because we're talking about the Parallel transported vector and the only reason we can do it is because we understand what this is and the only reason we know what that is is because we've defined it to be: $-\Gamma^{\mu}_{\alpha\beta}X^{\alpha}(P)\delta^{\beta}_{p}$ like that. This is what this little term is $\bar{\delta}X^{\mu}(P)$ and that sign convention you'll see in a moment why we use that sign convention.

We know what the parallel vector is so we can assume that we can calculate that once we know what this is once we know what the connection is on the manifold and then we can subtract these two and we can find a difference between the value of the Vector field at Q and the value of the Vector field at P by taking the value of the vector P , transporting it over here at Q , Parallel transporting it which is entirely driven by this idea, this connection coefficients $\Gamma^{\mu}_{\alpha\beta}$, the connection and that will provide us a vector at Q that we can compare with the new value of the Vector field there and it's important not to confuse this point.

Parallel transporting the value of Vector field at one point, Parallel transporting it to another point does not give you the value of the Vector field at this point Q , that's some independent number at some independent thing the Vector field is something that's could be arbitrarily established on space-time as long as it's smooth everything has to be smooth and differentiable because in being a differentiable manifold that is a well-defined idea, being smooth and differentiable but this parallel transport process is a totally different process than evaluating the Vector field at P in a Vector field at Q . Parallel transporting the vector at P to Q is a process that depends on the manifold property, calculating the value of the Vector field at P and calculating the value of the Vector field at Q only depends on the nature of the Vector field itself because it is a function of the coordinates so these are two independent processes. Given that this is infinitesimal distance, this path question doesn't matter, this path question becomes relevant when you have these finite distances involved but when you have infinitesimal distances this path it doesn't matter what the path is it's all going to limit to zero so now once you can do that and I can compare this Vector field of Q with this thing I can actually create something that's the derivative of this object and let's quickly see how to do that.

I'm going to start by writing down the difference between these two vectors the value of the Vector field at Q and from that I'm going to subtract the value of the Parallel transported Vector field:

$$X^{\mu}(Q)-\left[X^{\mu}(P)+\bar{\delta}\, X^{\mu}(P)\right] \tag{44}$$

That's the difference between these two and this distance between them is this δ_p^{α} 4-component little differential in coordinates so I'm going to take that difference and I'm going to divide it by this δ_p^{α} differential element and then I'm gonna take the limit so we're gonna do this in case this is regular calculus now, as $\delta_p^{\alpha} \rightarrow 0$:

$$\lim_{\delta_p^{\alpha} \rightarrow 0} \frac{X^{\mu}(Q)-\left[X^{\mu}(P)+\bar{\delta}\, X^{\mu}(P)\right]}{\delta_p^{\alpha}} \tag{45}$$

That looks just like a derivative but remember it's a derivative where I have some left over indices, if you look at it, I have an upper index here, this μ and I'm gonna have a lower index in the denominator this α so whatever this object is, it's going to have an upper index and a lower index so we could say would have a μ and an α so I'll call that it's equal to some object with a μ and an α , an upper and a lower index. Now the question is, is it a tensor object and the answer is, yes it is because this the first part of the numerator is a vector at Q and the second part is also a vector in Q because this is a Parallel transport, the first is a $(0,1)$ tensor at Q and the second is also a $(0,1)$ tensor at Q , the difference of course is another $(0,1)$ tensor at Q and the denominator is are real numbers so we're subtracting two vectors and dividing by real number so we know that this object is going to be a tensor like object in fact it's going to be a $(1,1)$ tensor object so let's work this a little bit so we're still going to take the limit as $\delta_p^{\alpha} \rightarrow 0$:

$$\lim_{\delta_p^{\alpha} \rightarrow 0} \frac{X^{\mu}(P+\delta_p)-X^{\mu}(P)-\left[-\Gamma^{\mu}_{\beta\alpha} X^{\beta} \delta_p^{\alpha}\right]}{\delta_p^{\alpha}} \tag{46}$$

Now look how I did this, in this term $P+\delta_p$ there's no indices on this one but here δ_p^{α} there's an index and that's very relevant because I'm talking about the point Q and the point Q is this infinitesimal distance away from P but what I'm interested in is the Vector field evaluated at the coordinates of P and the Vector field evaluated at the coordinates of Q that's what this is saying without the indices I'm at the coordinates of P plus the coordinates at Q so this first term is really:

$$X^{\mu}(P+\delta_p)=X^{\mu}(P^0+\delta_p^0, P^1+\delta_p^1, P^2+\delta_p^2, P^3+\delta_p^3) \tag{47}$$

That's what this means and that's a shorthand for that and you'll notice because it's in the argument and I'm trying to evaluate it at a point I'm just gonna write it without space-time indices. That's my job, my job is to interpret all of this notation for you so you guys can understand. I'm going to split this up:

$$\lim_{\delta_p^{\alpha}\to 0}\frac{X^{\mu}(P+\delta_p)-X^{\mu}(P)}{\delta_p^{\alpha}}+\frac{\Gamma^{\mu}_{\beta\alpha}X^{\beta}\delta_p^{\alpha}}{\delta_p^{\alpha}} \tag{48}$$

Now you see why we had that convention minus sign, that convention minus sign was here because ultimately what we want it be nicer to have a plus sign in the end so we define it with this little minus sign in it. Now this, in the second term will cancel. This is the limit as $\delta_p^{\alpha}\to 0$ so in principle this is 4 equations so I can do one for δ_p^0 and if that's the case, if I do that then you can see how it immediately will cancel out so let me let me just execute this:

$$\lim_{\delta_p^{\alpha}\to 0}\frac{X^{\mu}(P+\delta_p)-X^{\mu}(P)}{\delta_p^{\alpha}}+\Gamma^{\mu}_{\beta\alpha}X^{\beta} \tag{49}$$

This will ultimately be an object with two indices: upper μ and lower α . The first term in (49) is the ordinary partial derivative of X^{μ} with respect to a small displacement from P which is essentially x^{μ}

$$\lim_{\delta_p^{\alpha}\to 0}\frac{X^{\mu}(P+\delta_p)-X^{\mu}(P)}{\delta_p^{\alpha}}=\partial_{\alpha}X^{\mu}(P) \tag{50}$$

That is just your ordinary partial derivative of a function and it's not a tensor but we're adding to that this term $\Gamma^{\mu}_{\beta\alpha}X^{\beta}$ and the whole sum is in fact a tensor and this object we will give a special symbol to this symbol is $X^{\mu}_{\|\alpha}$ where that double bar basically means we're doing this process (51) and it's a $(1,1)$ tensor so the components are:

$$X^{\mu}_{\|\alpha}=\partial_{\alpha}X^{\mu}+\Gamma^{\mu}_{\beta\alpha}X^{\beta} \tag{51}$$

Another way to write it:

$$X^{\mu}_{\|\alpha}\,\partial_{\mu}\otimes dx^{\alpha}=\left[\partial_{\alpha}X^{\mu}+\Gamma^{\mu}_{\beta\alpha}X^{\beta}\right]\partial_{\mu}\otimes dx^{\alpha} \tag{52}$$

That is the Covariant derivative of a Vector field and it all depended on our ability to transport this vector $X^{\mu}(P)$, Parallel transport this vector to Q and then take the difference, everything hinges on that ability. Now notice when you Parallel transport an infinitesimal amount then the answer is simple, you know what this parallel vector is you, just add this little piece $-\Gamma^{\mu}_{\alpha\beta}X^{\alpha}(P)\delta_p^{\beta}$ because it's this infinitesimal amount but we've eliminated the need to actually execute this process of Parallel transporting in calculating this guy because I just did it, I took the limit. the limit goes to zero this is all that's left (51), the Covariant derivative of a Vector field is the partial derivative of that Vector field with plus this extra piece $\Gamma^{\mu}_{\alpha\beta}X^{\beta}$ that is proportional to the Vector field and you're left with an object that is a $(1,1)$ tensor and other question is what is that object, what can you do with it, in what sense is it a derivative? Well we know in what sense it's a derivative, it's literally a derivative, it's just like any function, it's the function at one point minus the function at another point with an infinitesimal difference driven to a limit, that's the classical simple definition of a derivative so it clearly is a derivative, the fact that it's a $(1,1)$ tensor is a bit puzzling perhaps and as we get to use these things more and more we'll see why that matters in fact we're going to talk a bit more about its symbolic representation in a moment before we close this abstract lecture.

Actually I thought about it it's already gotten longer than I thought so we will close the lecture now and what we've done is I've gone back and I've explained to you why $\Gamma^{\mu}_{\nu\lambda}$ is so important and how the whole concept of taking a derivative inside a manifold of tensor and Vector fields, we demonstrated for Vector fields but it generalizes to tensor fields and to co-Vector fields and I've explained you why this guy is so important, this guy generates parallelism and you'll notice in that discussion of the Covariant derivative and our discussion of $X^{\mu}_{\|\alpha}$ an the difficulty generating these components I never referred to the metric because we didn't have to, all you need to do is refer to these coefficients $\Gamma^{\mu}_{\nu\lambda}$ and eventually what we're gonna see is the significance of those coefficients the ability to take real derivatives in the physical world will ultimately depend on the metric because the metric will provide us with those coefficients and that's the significance of the metric connection and we'll get there so in our next lesson we're going to talk a little bit more about the different notations for $X^{\mu}_{\|\alpha}$ and the different ways that it is expressed and used in a few different books and we'll start getting ready to actually express and explain where this expression (**1.3.3**) comes from, see you next time.

Lesson 10: CFREE notation and the covariant derivative

We are going to proceed moving down our study of this catalog of space times I kind of like this idea of using this catalog of space times as sort of our foundation because we're not going to follow a literal textbook but we're gonna go through this and if you can understand this catalog of space times you'll be a pretty good shape to understand General Relativity and the part of the book that we're working on right now is at the very beginning it's this chapter called **1.3 Basic objects of a metric** and the idea here is that once we have decided our fundamental axioms of General Relativity where we decided that space-time is modeled with a manifold which is 4D with a Lorentzian metric which means a metric of signature, in our case -2 , it could be $+2$ but we're going with a -2 convention which under the basic assertion of the Einstein equivalence principle at every given place in space-time we can find a coordinate system where the metric is:

$$g_{\mu\nu}=\begin{pmatrix}-1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \tag{53}$$

Furthermore the metric coefficients have a 0 first partial derivative meaning $g_{\mu\nu}$ the first partial derivative equals 0 which is the same as saying the partial derivative $\partial_\delta g_{\mu\nu}=0$ which is the same thing as saying $\partial/\partial x^\delta g_{\mu\nu}=0$. Once we've established that our space-time is a manifold that has this ability has this property at any point there is a coordinate system that is like this (53), that's the Einstein equivalence principle weight into the manifold then in order to proceed in to produce physics we have to understand derivatives and that means we have to have some kind of connection which we've introduced through this notation $\Gamma^{\mu}_{\nu\lambda}$. Look how they do their notation for this, look at (**1.3.3**) at the notation for the Christoffel symbol of the second kind which we have called the connection. Notice that the μ is right above the ν , they did not write it $\Gamma^{\mu}{}_{\nu\lambda}$ with some kind of space right there, many many textbooks do, in fact most textbooks do, but the catalog of space-time I think correctly has made the choice not to do that, now the reason that they don't do that is because this object is not a Tensor, the sequence of those indices is irrelevant it is simply there, we can design this object any way we want and in fact it has been designed many ways there's a symbol where you have $\begin{Bmatrix}\mu\\ \nu\,\lambda\end{Bmatrix}$. It's literally a symbol like just drawn literally this way: two braces and three numbers and those numbers are indices and that is the value of one of these Christoffel symbols and there's alternative ways of doing that is $\{\mu,\nu\,\lambda\}$. It's just a choice of how you want to make the notation look for these connection coefficients The point is that we're choosing how to design this. When we do the Tensors, when we have like g as a Tensor and in this case $g^{\mu\rho}$, because we know it's a Tensor we have established its property says that we know that we're gonna have a Tensor product like this $g^{\mu\rho}\partial_\mu\otimes\partial_\rho$ next to it and this is the component we need. We know that this component implies this hidden implied Tensor product and these are the coordinate basis vectors of our manifold so when you see an index on something that you know to be a Tensor it drags in a lot of implied information in this case $\Gamma^{\mu}_{\nu\lambda}$ when you see an index it's not so much, there's no structure there's nothing that you add on to the back end of a connection coefficient in fact if you look at this $g^{\mu\rho}$, it should be obvious, this contracts with the ρ and what are you left with? You have this up that matches there and then you have two down indices which match here you would think well that sounds like a Tensor but you have these partial derivatives that are not covariant derivative so we just did that old darn lesson a partial derivative that's not covariant derivative it doesn't mean anything it just is not tensorial so you know that these objects in here are not Tensors just because the derivatives that they're dealing with are our regular partial derivatives.

Immediately you know this isn't a Tensor and in fact this whole thing has to be evaluated at a point and therefore you know that it is a number now also the other thing to remember this catalog is that everything is a field, everything is a Tensor field. We stopped writing Tensor fields but $g_{\rho\nu}$ is a function of space-time $g_{\rho\lambda}$ is a function of space-time and likewise you take the first derivative of a function of space-time you wouldn't be able to take a derivative, unless it was a function of space-time you wouldn't be able to take a derivative with a space-time index but that also means that this $\Gamma^{\mu}_{\nu\lambda}$ too is a function of space-time and this could easily be written $\Gamma^{\mu}_{\nu\lambda}(x)$ a function of space-time and so these are the things we have to keep in mind as we go through what I'm about to do next which is we're headed now in the direction of deriving this relationship here (**1.3.1**) which involves a symbol that we haven't seen before this symbol here $\Gamma_{\nu\lambda\mu}$ is not the same as that symbol there $\Gamma^{\mu}_{\nu\lambda}$ obviously they're very closely related and we're going to derive maybe perhaps more importantly this relationship (**1.3.3**) and as we said in the previous lectures we are now showing that the connection, this all-important notion of parallelism, can be determined purely by knowing the metric, the metric and the connection are not independent in the theory of General Relativity, in mathematics of manifolds they are independent, you can choose any metric you want on a manifold and then you can choose any connection you want and I could write the manifold $(S, \Gamma^{\kappa}_{\nu\lambda}, g_{\alpha\beta})$, that's the set and then I could say S has some connection $\Gamma^{\kappa}_{\nu\lambda}$ in which I would maybe write in component notation like that and it has a metric $g_{\alpha\beta}$. There's three things that it has and that's fine you can do that mathematically but in physics these two are bound together and the structure of space-time which is really defined by the connection through the notion of parallelism that structure of the manifold is determined by the metric on the manifold and so these things get bound together.

We're gonna start marching towards understanding this but I'm gonna kind of kill two birds with one stone, that is, I'm going to be exploring alternative I don't want to say notations but an alternative approach to the mathematics of General Relativity that doesn't involve components it involves the general abstract Tensor notation that is a little bit more popular in modern treatments and actually it shows up extensively and Misner, Thorne and Wheeler (MTW) and that's like 1972 so it's hard to call that modern but when you talk about Science and Math, if it's after 1920, it's probably modern and there's stuff in 1920 that is considered modern but MTW heavily leans on a branch of notation that doesn't involve components at all it's the pure Tensor notation the abstract Tensor notation and so I'm gonna call these two different types, I'm gonna give them a name so I can refer interchangeably, I refer to the two types of notation. The first is gonna be called the comp notation and comp stands for components and the second is gonna be the CFREE notation which means I'm going to call that the “component free” notation and we are going to constantly be going between these two so that you have a lot of flexibility to go through any different textbook that you ever see out there but be advised that the catalog itself entirely operates in the comp system.

Likewise by the way on this project of the basic objects of a metric we're going to show that the connections, this connection here in particular (**1.3.3**) is determined by the metric but what we also see is that there are other Tensors: the Riemann Tensor (**1.3.5**) is all determined by the metric and how do I know that well the Riemann Tensor is defined by the connection and derivatives of the connections partial derivatives are the connection but the connection is defined by the metric therefore the Riemann Tensor is defined entirely by the metric likewise the Ricci Tensor (**1.3.9**) is defined entirely by the Riemann Tensor and the metric therefore this is entirely defined by the metric and likewise the scalars and all these other things everything once you have the metric all these other things fall into place, it's not trivial, $\Gamma^{\mu}_{\nu\sigma,\rho}$ is an extensive mess of a calculation, if you tried to put this in terms of the metric because each one of these you would have to substitute in this take a derivative which would now involve second derivatives of the metric, in fact that's this is you can just without even studying what these objects are, we haven't even talked about what the Riemann Tensor is, but you can see a lot right from the catalog itself, you can see right away that the connection involves the first partial derivatives of the metric and you can see that the Riemann Tensor evidently is a combination of the connection but this is a Tensor so it shows you that this combination even though $\Gamma^{\mu}_{\nu\sigma,\rho}$ is not a Tensor, this combination (**1.3.5**) is in fact a Tensor and what that always means is that the non tensorial part that means if you transform this from one coordinate system to another it wouldn't transform like a Tensor but the part that's left over that makes it non tensorial, all of it must cancel and in fact if you multiply these two together $\Gamma^{\mu}_{\rho\lambda}\Gamma^{\lambda}_{\nu\sigma}$ this must be a Tensor because you're adding terms so each term must be a Tensor so this is a beautiful cancellation here what's left over in fact is a Tensor but these guys here involve first derivatives of the connection which they must involve second derivatives of the metric.

I've already explained that based on the Einstein equivalence principle there's always a coordinate system where the first derivative vanishes at a certain point, remember this is a function of space-time so when I say vanishes at a certain point I mean there is a point in space-time, $g_{\alpha\beta,\delta}(x)$ is a function of space-time and when if I put that function at a certain point P it equals zero. Sometimes what we do is you write this equal sign where you would put a dot up here and that dot means it equals zero in a specific coordinate system, without that dot I'm saying that the first derivative of the metric is zero everywhere. If I just wrote $g_{\mu\nu,\delta}=0$ that's telling me that in all coordinate systems the metric is zero and in this case it would be at every point in space-time. Just by looking at this page alone is teaching us something because what I can say here is that even if I use a coordinate system that's flat at a specific point, if I evaluate this Tensor at a specific point this is a Tensor component that is a function of space-time, I put in that space-time point and I can evaluate this as a number or I can evaluate for example $R^{0}{}_{112}$, that's a component of the Riemann Tensor I could evaluate at the point x , well there's no guarantee that all of these will be zero even in this special coordinate system where the first derivative of the metric is zero because now I'm taking the second derivative of the metric and there is no guarantee by the Einstein equivalence principle that the second derivative of the metric would look like $g_{\rho\nu,\lambda\gamma}$ or something like that, there's no certainty that that will be zero so this Riemann Tensor may not even, in space time in a coordinate system where you've got a Lorentzian Cartesian system at a certain point you still may have a Riemann Tensor and the purpose of that is to say that you can put yourself into this inertial coordinate system that is flat but flatness does not mean that the curvature is gone, flatness just means that the metric at that point is this classic Lorentzian metric (53) and any divergence, as you move away from that point, is a second order dependence and $\boldsymbol{R}^{\mu}{}_{\nu\rho\sigma}$ is the second order dependence and it's another way of talking about tidal forces and we'll get into all of that later.

Now I want to start and working with this abstract notation a little bit and slowly start getting here it's gonna take a lecture or two, I don't think I'm gonna finish it today. Let's begin, we will start by having a look at, we're gonna have our space-time with our coordinate system but I'm going to now kind of not ever reference the coordinate system normally we talk about (X^0, X^1, X^2, X^3) and our basis vectors are ∂_μ and dx^μ like this, now we're not going to reference this at all, we're going to now talk about any coordinate system, we're not going to specify a basis and we're going to speak in the total abstract. When we talked about our connections we talked about $\Gamma^{\mu}_{\nu\lambda}$, these are in fact space-time indices but not on a Tensor but there's still space-time indices and they're meant to be contracted with other Tensors so for example a vector X^ν could contract on this ν value and this one is a vector, it is a $(1,0)$ Tensor so we no longer gonna have access to these indices because we're no longer talking about any specific coordinate system, we are now going to completely sort of erase the specificity in our mind and we have to reconstruct all of this in the pure extract sense.

How do we do this? Well if this is our space-time S I'm now gonna say instead of S with a connection $\Gamma^{\mu}_{\nu\lambda}$ and all that that implies, I'm still sticking with S as a manifold which means coordinate systems do in fact exist, arbitrary coordinate systems exist, those mappings that can take the manifold into $\mathbb{R}^4$ all of those exist but we're not going to specify one but the set still exists and now I'm going to define my connection with this symbol ∇ and if I had a book it would be $\boldsymbol{\nabla}$, that will be the abstract notion for the connection. No components, it's just our symbol for the connection and we treat the connection as an operation and as an operation the MTW way of presenting it is as a machine that takes three objects and then it returns, if it's given those three objects, it returns a real number so the middle object that it takes is some sort of Vector field so I might say I'll call it a Vector field $\vec{A}(x)$. We have no coordinates so I don't want you to confuse this with the components of the Vector field but if I have no indices on it how can you make that confusion I'll try to bold it out but I'm going to put a little arrow on it. The way you know it's a Vector field is I'm explicitly showing its space-time dependence but then on top of that it's also going to have another $\vec{B}$ but $\vec{B}$ is not a Vector field, it's a vector located at a certain point so $\vec{B}$ at some point P . Now $\vec{B}$ could be a Vector field but all we are interested in is evaluating this this will give us a real number if we fill in all the slots but we have to do it at a point P .

The middle slot actually takes a Vector field though so even though I want a vector located at P and I want a Vector field that covers the part of the manifold that includes P , I need to separate vectors: one is just at a point and the other is a whole field and then in the first spot I have some kind of of form this is going to be a 1-form ω and again it's just a 1-form that's located at P and if I get all of this down then I have a real number:

$$\nabla\left(\omega(P),\vec{A}(x),\vec{B}(P)\right)\in\mathbb{R} \tag{54}$$

If we were on our space-time S and I guess I have to draw the space-time like this now because I'm not going into a specific coordinate system and I pick out a point P then this machine, this ∇ , exists at this point P and then if there's another point Q , Q has its own ∇ , it's sort of a ∇ field and it's got these three slots (54). The key that I want to point out is the middle slot has to be a Vector field that makes sense on a neighborhood around P because this is ultimately the guy we're going to take the derivative of which is $\vec{A}(x)$ so it has to exist all the way around P but I emphasize that $\vec{B}$ only has to exist at P and ω has to exist at P . If I write $\vec{B}(P)$ I'm already kind of implying that $\vec{B}$ is a Vector field and I'm just evaluating it at the point P . I don't really mean that, I just want it to be a single vector at P , maybe I should write it $\vec{B}_P$ and that implies it's just a single vector in the tangent space at P , in the tangent space $T_P(S)$ (sometimes you see it in books with parentheses around, this is meant to be a set of vector space called the tangent space of P which I presume you already fully understand.

In order to get this machine to evaluate to a real number I need all these things, I need a neighborhood around P which has a Vector field associated with it, I need ω_P and I need a $\vec{B}_P$. Normally in a textbook these would be boldface like I said. If I have (54) at every point then that is a connection on the manifold as long as (54) exists at every single point so it's like a field in itself, it has the sense of being a Tensor, a Tensor has slots for vectors and co vectors and here's there is a vector and a co vector so that's the same but this middle slot being a field, that's different so this machine is different. That is the connection, that's the abstract understanding of the connection, it's this map that takes these three things and gives a real number. Now just like a Tensor I can consider this machine and leave the first slot blank put in the field and put in the vector at the point P . I'm going to suppress the P for now, we all understand that that's just a vector at this single point P :

$$\nabla\left(, \vec{A}(x), \vec{B}\right) \tag{55}$$

Now this guy is going to be an object that will give you a real number but it's waiting for a 1-form. The next abstract step here is we need to say, how do you get a real number out of a 1-form? Well you can take the dual space mapping of the 1-form with some vector and that will be the real number so I need to take these two guys (2nd and 3rd slot) and create some sort of vector, a simple straight up vector, if I could do that then I know immediately how I'd get my real number I just take that vector and get a dual space map from that vector. I am going to then say let's imagine that process is possible and I'm gonna give that process a symbol and that symbol is going to be ∇ , it's going to be the connection ∇ , I'm gonna put $\vec{B}$ as a subscript, as a bold subscript, it would be bold in a book and then I'm going to put the field a right here and I'm going to drop this x i.e. $\nabla_{\vec{B}}\vec{A}$. You have to understand when you see this that we're talking about a vector located at a certain point and $\vec{A}$ is a Vector field.

$\vec{B}$ could be a Vector field but all that matters is its value at a certain point P so even if $\vec{B}$ was a Vector field that'd be fine but the field nature of it would be irrelevant only the value at the point P matters as far as $\vec{A}$ goes the field nature is important because ultimately there's derivatives involved hence we're talking about the change of $\vec{A}$ as you move about the manifold so it has to be a field otherwise you can't really talk about the change at all, but this symbol here is supposed to be a member of the tangent space at the point P . $\nabla_{\vec{B}}\vec{A}$ is a vector so in principle if I was to switch this to the component notation and I called this $\vec{X}$ then I would be able to write that $\vec{X}=X^{\alpha}\partial_{\alpha}$. If I went to a coordinate basis, if I introduced a coordinate system on the manifold then I could take this abstract $\vec{X}$, I could write it this way $\vec{X}=X^{\alpha}\partial_{\alpha}$ and it would be just a regular vector member of the tangent space, now in the abstract form, I don't do that, I just leave it as this symbol and I know that this is a vector and we call that vector well a lot of people call it different things, the Covariant derivative of $\vec{A}$ in the direction $\vec{B}$, that's one of the things you could call that, the directional covariant derivative is another way. Once I have that, the way I create my real number is I put it here $\left\langle \omega, \nabla_{\vec{B}}\vec{A}\right\rangle$ and voila, it's a real number.

In true abstract form I haven't come up with a prescription of how to calculate this but I'm now proposing that I can you know that if we could we could use it in this fashion to get the real number and if we could create a vector out of a Vector field and this vector that's appropriate then we would have a fully understood connection on the manifold. Now we need to talk about $\nabla_{\vec{B}}\vec{A}$, we've boiled down the conversation to this guy so let's talk about that. We are working on understanding this expression $\nabla_{\vec{B}}\vec{A}$ where $\vec{A}$ is a function of space-time and, like I said, $\vec{B}$ could be a function on the space-time as well but we're only interested in $\vec{B}$ at a certain point P and we're going to interpret this as how rapidly is the Vector field $\vec{A}$ a changing along a curve, so there's some curve involved, along a curve, will call that curve $\gamma(\tau)$ where τ is the parameter for the curve, along the curve where the curve has a tangent vector at the point P where that tangent vector is $\vec{B}$ so along this curve, this curve has a tangent vector at P and that tangent vector is $\vec{B}$, we are asking how is $\vec{A}$, the field is changing and you might say, well here's $\vec{A}(P)$ and here's the field $\vec{A}(Q)$ and the question is how much is that Vector field changing along this curve and the tangent at that curve is $\vec{B}$?

Obviously P and Q , they're pretty far apart because there's not even linear approximation on $\vec{B}$ but the idea is $\vec{B}$ is tangent to some curve and it reflects the magnitude that that curve is changing, the value of how fast the point on the manifold is changing with respect to the parameter τ , that's gonna be the magnitude of $\vec{B}$ and the direction of $\vec{B}$ somehow indicates how the new coordinates at $\tau+\delta\tau$

are going to be related to the old coordinates of P . Remember, P in principle, since you're on a manifold you know it has coordinates $P \rightarrow (x^0, x^1, x^2, x^3)$ and a nearby point Q it has coordinates $Q \rightarrow (x^0+\delta x^0, x^1+\delta x^1, x^2+\delta x^2, x^3+\delta x^3)$ and it's these δx that are really fundamental to the concept of how fast $\vec{A}$ is changing because $\vec{A}(P)=\vec{A}(x^0, x^1, x^2, x^3)$ remember it's a manifold so I can always do that and $\vec{A}(Q)=\vec{A}(x^0+\delta x^0, x^1+\delta x^1, x^2+\delta x^2, x^3+\delta x^3)$. I'm violating this notion of the totally abstract notation because here we actually have to go into a coordinate system so we're going to avoid doing that but the point is that this object $\nabla_{\vec{B}}\vec{A}$, whatever it is, is supposed to discuss how fast the vectors field is changing at this point P but it's along a curve whose tangent vector is $\vec{B}$ at an entire curve whatever this curve is, is completely distilled for the purposes of this analysis, of this rate of change analysis, to its tangent vector.

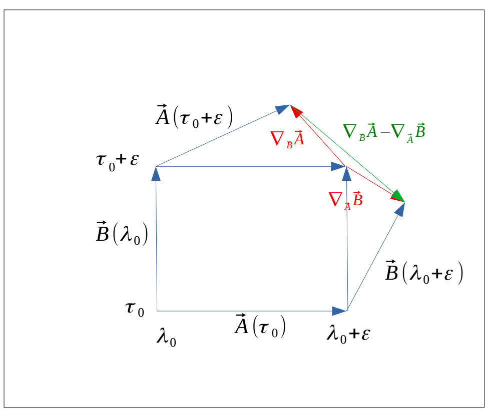
The curve actually kind of disappears and this is an important and confusing point to most students I think, when I say that what I mean is it was confusing to me and my friends and some graduate students who I've worked with over the years but when I'm trying to calculate something like this you can put any vector down here in this subscript you could put $\nabla_{\vec{X}}$, $\nabla_{\vec{Y}}$ whatever the vector is and the relevant curve will change, the curve totally disappears except for the tangent vector at that point. There are many different curves that have the same tangent vector at a point but it all doesn't matter because you're only going to this limit where just the tangent vector is going to be the only thing that's important to calculate this difference.

If we now look at just sort of a classic textbook analysis of this, if you want to follow along MTW, Page 249, Section B, I'm just gonna duplicate that here because it's actually a pretty good way of understanding what this symbol is $\nabla_{\vec{B}}\vec{A}$ and the way he wants us to understand it is in this fashion we have $\vec{A}$ and that is a Vector field and the parameter of $\vec{A}(\tau_0)$ that is the value of some vector at τ_0 , now he's doing this thing where he is putting the vector, he's attaching the vector to a point so this point here would be $\gamma(\tau_0)$ that's what that point would be, now in his analysis he just puts it as τ_0 because he's now saying look you know that if you have a manifold and you have a curve on that manifold you have a real line with τ and at some point τ_0 , you get $\gamma(\tau_0)$ at some point P in the manifold, so why don't we skip all this and just call P as τ_0 ? We know we're on a curve we know the curve has a parameter let's just call it τ_0 so that's what he does in his analysis, well I guess MTW are three different people's so that's what they do so they just call the point τ_0 .

I'd be inclined actually to go all the way and call $\vec{A}(\tau_0)$ as $\vec{A}(\gamma(\tau_0))$ because remember $\vec{A}$ is an element of the tangent space of P so I need to get a $P=\gamma(\tau_0)\in S$. I like going all the way with this and putting this whole thing in there because this is what causes a lot of confusion and struggle is what are all these symbols actually meaning. They also put that vector on the point which I kind of have been pointing out I don't like that my inclination would you take this and create a little box and then create a vector in that box but you're just gonna have to do that in your head because that's not how MTW does it. We wanted to understand the rate of change of this thing so we're kind of looking for this notion of $d\vec{A}/d\tau$ as this parameter changes as we move along this curve, as we start from P and we move along this curve τ changes and $\vec{A}$ is a Vector field so it's gonna have a different value at different points depending on how the Vector field was established on the manifold as long as it's smooth, that's the only real requirement.

With that in mind we now need a nearby point and that nearby point will be over here and the Vector field will have a value at that point too and the way they draw it is like this $\vec{A}(\tau_0+\varepsilon)$ (some small amount) and the vector that connects those two points is going to actually be a vector that we are going to call $\vec{B}$ that's our vector $\vec{B}$. Now the idea is here, we have this is part of that awkward problem we have two points in the manifold this is what we're calling well this is now $\gamma(\tau_0+\epsilon)$, that's that point so we would call the first point P and say the second point Q and they're both along this curve. Well the point is that what you're now saying is that this point Q on the curve is so close to this point P that we can literally use the components of this vector to change the coordinates of this point P into the coordinates of that point Q . We're taking a limiting procedure where Q and P are so close together that $\vec{B}$ can be treated as a linear change in the coordinates themselves and that is that's fine actually in one sense because of the Einstein equivalence principle which says that there is a coordinate system out there where this is all flat space-time and in flat space-time you can actually embed these vectors inside the manifold so you could think about it that way. The point is that this small difference is completely captured by just the tangent vector on the curve.

Now the idea is you parallel transport this guy back to the point P and when you're this close you can parallel transport it in a literal way you treat everything as sort of flat. This thing here $\nabla_{\vec{B}}\vec{A}$ that is simply the vector difference between the parallel transported vector at $\vec{A}(\tau_0+\varepsilon)$ so that's $\vec{A}^{\parallel}(\tau_0+\varepsilon)$ from Q to P . The difference between this vector here which is $\vec{A}(P)$ and the parallel transported version of $\vec{A}(Q)$ and I take that limit as $\epsilon\to 0$ and that guy is $\nabla_{\vec{B}}\vec{A}$ and that is how MTW likes to graphically describe this picture and it's important to also understand that you could do this in the reverse you could take this vector at P parallel transport it to Q and then the difference is $\nabla_{\vec{B}}\vec{A}$ on the limit when $\epsilon\to 0$ because as long as you can parallel transport it either way and so in this case the parallel transport sort of stretches the vector out a little bit but if that's important you can go either direction and in fact I'm gonna close this lecture with a little demonstration.



If you had two Vector fields $\vec{A}(x)$ and $\vec{B}(x)$ you can discover an important and interesting fact, you start with $\vec{A}(\tau_0)$ and you can Parallel transport to $\tau_0+\epsilon$ and then this is $\vec{A}(\tau_0+\varepsilon)$ then this vector here is $\nabla_{\vec{B}}\vec{A}$ but in the same way you can actually go in the other direction: you can take $\vec{B}$ because now $\vec{B}$ is a Vector field, we're using $\vec{B}$ as a tangent vector of some curve about this point τ_0 to make the change from the point that we've labeled just τ_0 and the spirit of MTW and transport it from τ_0 to $\tau_0+\epsilon$.

but you could also take $\vec{A}$ as the tangent to the curve and transport $\vec{B}$ over to some other point along some other curve, you would actually need to give this curve that's got τ as a parameter, now you have to imagine some arbitrary curve that $\vec{A}$ is covering with parameter λ so this point here is λ_0 and this is $\lambda_0+\varepsilon$ and now we have $\vec{B}(\lambda_0+\varepsilon)$ whereas this is $\vec{B}(\lambda_0)$. Now we do the same kind of arrangement we transport $\vec{B}$ from from λ_0 to $\lambda_0+\varepsilon$ and we get some vector that looks like this, that's the Parallel transported version because in this picture when we do our parallel transports it's nice to just keep them literally parallel just to indicate that those are the parallel vectors but when we do that we end up with this vector here which is now $\nabla_{\vec{A}}\vec{B}$ and then what we do is we realize the final step is that this vector here would be $\nabla_{\vec{B}}\vec{A}-\nabla_{\vec{A}}\vec{B}$ and what's interesting is that this is not zero and so the failure to close of this polygon has to do with the fact that parallel transport in different directions is just not equivalent and that's to show that to make that little demonstration is in MTW, Page 250 Section C so it's really the next box. We'll talk about this later when we talk about the Riemann curvature tensor but right now we've just gone through in this lecture the basic idea of this abstract symbol $\nabla_{\vec{B}}\vec{A}$ and what it's supposed to mean and the basic point was this idea it is the difference between the parallel transported vector and the Vector field value at a point being parallel transported to another point and the difference between that and the Vector fields value at that new point that difference is what we call the covariant derivative of $\vec{A}$ in the direction $\vec{B}$, that's what this symbol means and next time we're going to start abstractly tearing this thing apart and understanding it in detail.

[Lesson 11: CFREE to COMP conversion]

We're going to continue this lecture about this abstract idea of the Covariant derivative of the Vector field $\vec{A}$ with respect to $\vec{B}$ i.e. $\nabla_{\vec{B}}\vec{A}$. Now I'm gonna leave out the arrows or x dependency just like that, we understand $\vec{A}$ is a full Vector field and $\vec{B}$ is also a Vector field perhaps but its value at the point P is important, we'll get back to that in a moment so we're gonna sort of understand it all, we're gonna start taking this and speaking of it in an axiomatic way. We want it to be a derivative so we're gonna have to say: what if we make certain changes, what if I ask for the Covariant derivative of the Vector field $\vec{A}$ with respect to some real number f multiplied by $\vec{B}$ i.e. $\nabla_{\mathrm{f}\vec{B}}\vec{A}$. Now if I understand $\vec{B}$ to be a Vector field and I multiply it by a number a I'm just sort of multiplying this whole Vector field by a constant but what we're usually talking about is a function on the space-time $\mathrm{f}(x)$, a Vector field on the space time $\vec{B}$ and in this case what I'm particularly interested in is that function at $\mathrm{f}(P)$ and the Vector field at $\vec{B}(P)$ because as I've said several times all that matters is at a certain value, at a certain point in the space-time manifold so I'm curious about how this derivative might change if I multiplied the Vector field $\vec{B}$ by some arbitrary function on the space-time where we always assume of course a smooth function.

You could imagine that if you were doing regular calculus and your independent variable was x and you suddenly converted $x\rightarrow 3x$ you would imagine that the derivative, well let's say you had the function x^2 but now you're changing $x\rightarrow 3x$ so the derivative of x^2 is $2x$ but the derivative of $(3x)^2$ would now be $6x$ so we're interested in this idea that it's should be multiplicative because that's the independent variable in a sense, that's the direction of change so we're going to embrace that and we're going to insist that:

$$\nabla_{\mathrm{f}\vec{B}}\vec{A}=\mathrm{f}\,\nabla_{\vec{B}}\vec{A} \qquad (56)$$

All you've done is scale $\vec{B}$, it is going to be the scale of the directional derivative of $\vec{A}$ in the direction of $\vec{B}$ without the scaling times the scale factor. Remember this guy here $\nabla_{\vec{B}}\vec{A}\in T_P$ that's a vector an element of the tangent space at S , a tangent space of the point $P\in S$ and (56) is also a vector, an element of the tangent space at S and f is a real number and you're allowed to multiply vectors by real numbers as long as it's a real vector space, that's the whole point of a real vector space is you can multiply it by scalars and in this case the scalars are the real numbers so that's what (56) means.

(56) is an axiomatic statement, this is a demand we make on the nature of this object so that's one thing that's how you scale that but now the question is what about scaling $\vec{A}$ with the same kind of function what if I did $\nabla_{\vec{B}}(\mathrm{f}\vec{A})$ where I'm now taking the Directional derivative of the Vector field $\mathrm{f}\vec{A}$ in the direction of $\vec{B}$. Now it's the same thing, $\vec{A}$ is a function of x , f is a function of x so this whole thing is some new Vector field $\vec{X}$ which is a function of x . What about this guy $\nabla_{\vec{B}}(\mathrm{f}\vec{A})$ well we want that to behave the same way you would have the [Product rule] behavior:

$$\nabla_{\vec{B}}(\mathrm{f}\vec{A})=\left(\nabla_{\vec{B}}\mathrm{f}\right)\vec{A}+\mathrm{f}\,\nabla_{\vec{B}}\vec{A} \qquad (57)$$

These two rules (56) and (57) are telling us how things operate but we need to understand $\nabla_{\vec{B}}\mathrm{f}$ and what would be if we imagine our manifold, we have the manifold S we have the point P and we have some function $\mathrm{f}(x)$ on the manifold and $\mathrm{f}(x)$ doesn't have vector components $\mathrm{f}(x)$ is just a number at every point in the manifold and it's derivative in the direction of some tangent vector along some curve which would be $\vec{B}$ in this case, that is simply going to be the directional derivative of the function $\mathrm{f}(x)$ with respect to its coordinate system. Now we already know that all vectors $\vec{B}$ that are elements of the tangent space $T_P(x)$, they are differential operators so we can now talk about the operation of $\vec{B}$ on $\mathrm{f}(x)$ and that will give us another function on the manifold: basically a partial derivative of $\mathrm{f}(x)$ on the manifold S , that would be in the CFREE notation, we would write it this way $\vec{B}(\mathrm{f})$ or we would just write it $\vec{B}\,\mathrm{f}$ and it's understood that we're taking $\vec{B}$ and applying it to the function $\mathrm{f}(x)$. to get a new function. In the comp notation we would write $\vec{B}$ is simply $B^{\mu}\partial_{\mu}$ so this guy is:

$$\left(B^{\mu}\partial_{\mu}\right)\mathrm{f}\rightarrow B^0\partial_0\mathrm{f}+B^1\partial_1\mathrm{f}+B^2\partial_2\mathrm{f}+B^3\partial_3\mathrm{f}=B^0\frac{\partial\,\mathrm{f}}{\partial x^0}+B^1\frac{\partial\,\mathrm{f}}{\partial x^1}+B^2\frac{\partial\,\mathrm{f}}{\partial x^2}+B^3\frac{\partial\,\mathrm{f}}{\partial x^3} \tag{58}$$

This is simply the partial derivative of $\mathrm{f}(x^0,x^1,x^2,x^3)$ with respect to the variable x^i of which it is a function because it is a function on the space-time and here I've broken the space-time into a specific coordinate system and this derivative is easy to take and $\vec{B}$ is literally defined by this process on some arbitrary function, in this case we would use it on f so $\nabla_{\vec{B}}\mathrm{f}$ is very easy to understand. I will write:

$$\nabla_{\vec{B}}(\mathrm{f}\,\vec{A})=\vec{B}(\mathrm{f})\vec{A}+\mathrm{f}\,\nabla_{\vec{B}}\vec{A} \tag{59}$$

Notice how everything works here, the second term has $\mathrm{f}\in C^{\infty}(S)$ which is a real function times a vector $\nabla_{\vec{B}}\vec{A}\in T_P(x)$, the second term is a differential operator acting on a function which returns another function $\vec{B}(\mathrm{f})\in C^{\infty}(S)$ and $\vec{A}\in T_P(x)$ is an element of the tangent space at s so the second term is also a vector so it's a vector plus a vector it's all done at the point P or at some point on the manifold and presumably can be done at any point on the manifold where all of our conditions are met. The conditions being that $\vec{A}$ is a Vector field defined on that point f is a function defined … $\vec{A}$ is a Vector field defined in the neighborhood of all the points, f is a function defined in the neighborhood of all the points, $\vec{B}$ is a Vector field that it has some value at the point, it doesn't have to be defined in the entire neighborhood of the point it does have to be defined along some curve but that's what a vector is, in the way we've done it a vector is always defined on some curve by its very nature.

These two rules, (56) and (59), are sort of our axiomatic foundations for the behavior of this idea of this Covariant derivative type operator. When I say these are the axiomatic rules I'm forgetting two critical ones which is the statements of linearity:

$$\nabla_{\vec{B}+\vec{C}}\vec{A}=\nabla_{\vec{B}}\vec{A}+\nabla_{\vec{C}}\vec{A} \tag{60}$$

Likewise:

$$\nabla_{\vec{B}}(\vec{A}+\vec{C})=\nabla_{\vec{B}}\vec{A}+\nabla_{\vec{B}}\vec{C} \tag{61}$$

These are the axiomatic rules that really define the full behavior of the Covariant derivative. You'll notice it's always in terms of itself. This does give us a hint on how to do our conversion from the CFREE which is all this notation, is coordinate free into the comp notation which is the coordinate base notation involving components and you can see a couple of examples that are worth diving into for example if I substitute $\vec{B}=B^{\mu}\partial_{\mu}$, we're in the full component notation, we've selected a coordinate system, we've now got the coordinate basis and the vectors are now all identified with differential operators, well let's go to the first rule $\nabla_{B^{\mu}\partial_{\mu}}\vec{A}$, well actually we need to go to the third rule first

$$\nabla_{B^{\mu}\partial_{\mu}}\vec{A}=\nabla_{B^{0}\partial_{0}+B^{1}\partial_{1}+B^{2}\partial_{2}+B^{3}\partial_{3}}\vec{A} \tag{62}$$

Now I'm going to use (56) and (60) because remember, this is a function of x , the components are a function on the manifold and the basis vectors are our partial derivatives taken at a point on a manifold so these basis vectors are all elements of $T_{X}(S)$ and then the components are are a function on the space-time that's why $\vec{B}$ is a Vector field, so f becomes B^{0} and B becomes ∂_{0} and I'll just do a combination:

$$\nabla_{B^{\mu}\partial_{\mu}}\vec{A}=B^{0}\nabla_{\partial_{0}}\vec{A}+B^{1}\nabla_{\partial_{1}}\vec{A}+B^{2}\nabla_{\partial_{2}}\vec{A}+B^{3}\nabla_{\partial_{3}}\vec{A} \tag{63}$$

Which in almost every book you'll see is going to write that as:

$$\nabla_{B^{\mu}\partial_{\mu}}\vec{A}=B^{0}\nabla_{0}\vec{A}+B^{1}\nabla_{1}\vec{A}+B^{2}\nabla_{2}\vec{A}+B^{3}\nabla_{3}\vec{A} \tag{64}$$

They're going to suppress that partial in the subscript and this generally books try to do things where you're not dealing with sub scripted subscripts. If you can get away without doing it. The bottom line is this, is going to be:

$$\nabla_{B^{\mu}\partial_{\mu}}\vec{A}=B^{\mu}\nabla_{\mu}\vec{A}=B^{\mu}A^{\nu}_{\parallel\mu}\vec{e}_{\nu} \tag{65}$$

That's a step towards this comp notation, it's not the full comp notation. The $\vec{e}_{\nu}$ is added to preserve the vectored nature. You might throw down the vector itself that sums over the μ but in true comp notation you don't do that, you'd leave this out that's the whole thing so this is actually the binding between the comp notation, the full comp notation and the abstract notation so I guess that's worth remembering.

$$\nabla_{B^{\mu}\partial_{\mu}}\vec{A}=B^{\mu}A^{\nu}_{\parallel\mu}\equiv\nabla_{\vec{B}}\vec{A}=B^{\mu}\left[A^{\nu}_{,\mu}+\Gamma^{\nu}_{\sigma\,\mu}A^{\sigma}\right] \tag{66}$$

That's how you're going to switch these things in your mind the covariant derivative of A contracted on B^{μ} is what this thing is. It's fair to do the full thing out, you could go that far right you could go ahead and substitute to this and then this would be the full comp notation although in the true comp notation you don't actually have the final vector part you just know it's a vector because this whole thing ends up being something with one contravariant index which of course is what we understand to be a vector. This is how you start applying those rules to convert, you apply these rules with these kinds of substitutions $\vec{B}=B^{\mu}\partial_{\mu}$ and you can convert from CFREE into comp notation pretty easily.

But you can't get all the way there. Remember when I made this last line I actually substituted for this $A^{\nu}_{\parallel\mu}$ based on our analysis that was done completely in the comp notation a couple lessons ago where we peeled apart how to describe this Covariant derivative of a Vector field so this guy here $\Gamma^{\nu}_{\sigma\,\mu}A^{\sigma}$ has to come from the field, it has to come from our knowledge of this analysis so the question is how do we get at this through this CFREE notation and the answer is to lean on this expression here (59) and how are we gonna do that? I'm gonna write down the Covariant derivative $\nabla_{\vec{B}}\vec{A}$ in the $\vec{B}$ direction and then I'm going to make the substitution:

$$\nabla_{\vec{B}}\vec{A} \Rightarrow \nabla_{B^{\mu}\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} \tag{67}$$

I'm going to use two rules to decompose this thing and so I'll begin with the easy one which will be:

$$\nabla_{B^{\mu}\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} = B^{\mu}\nabla_{\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} \tag{68}$$

I've kind of used a combination of (56) and (60) just like I did down here in (66) where I broke it all down and I used the linearity and then I re-collapse it back into this thing $B^{\mu}\nabla_{\mu} A$, that sort of what I did and you could also alternatively write this while getting rid of that $\vec{e}$:

$$\nabla_{B^{\mu}\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} = B^{\mu}\nabla_{\mu} A^{\nu}\vec{e}_{\nu} \tag{69}$$

I'll stick with the left hand side of (69) for now so then you use the Leibniz rule so now I have:

$$\nabla_{B^{\mu}\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} = B^{\mu}\left[\left(\nabla_{\vec{e}_{\mu}} A^{\nu}\right)\vec{e}_{\nu} + A^{\nu}\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu}\right] \tag{70}$$

Now let's do our analysis here, remember A^{ν} is a function, it's a function on space-time, we probably call it an infinitely differentiable function C^{∞} on the space-time. We have a definition for $\nabla_{\vec{e}_{\mu}} A^{\nu}$ this we know that the Covariant derivative of a function is just the vector operating on the function so now we kind of have to go back and think of this these vectors in terms of vector operators or of differential operators so we'll go with $\vec{e}_{\mu} \rightarrow \partial_{\mu}$ and we'll do that for all the vector so this becomes:

$$\nabla_{B^{\mu}\vec{e}_{\mu}} A^{\nu}\vec{e}_{\nu} = B^{\mu}\left[\left(\partial_{\mu} A^{\nu}\right)\vec{e}_{\nu} + A^{\nu}\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu}\right] \tag{71}$$

This guy $\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu}$, we can't really reduce any further so we're sort of stuck with that and in our system of abstract notation that's kind of where it the whole thing bottoms out so we need to be able to convert this thing into the comp notation and we just do that by sort of fiat definition, we're gonna say that that We know that $\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu}$ is a vector unlike this guy $\nabla_{\vec{e}_{\mu}} A^{\nu}$ which is a scalar function because it's the Covariant derivative of a function, it's not the Covariant derivative of a vector, the vector part was left out here $\vec{e}_{\nu}$ and it was the Leibniz rule that now we're dealing with the Covariant derivative of the vector part which is a vector and of course it's times a function so everything works out. You have a function times a vector which is another vector and you have a function times a vector which is another vector, you're summing two vectors using the addition rule of vectors and then you're again multiplying by functions and ultimately again this whole thing is presumably a vector. This vector $\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu}$ in component notation is going to be:

$$\nabla_{\vec{e}_{\mu}}\vec{e}_{\nu} = \Gamma^{\alpha}_{\cdot\cdot} \cdots \vec{e}_{\alpha} \tag{72}$$

i.e. it is some coefficient times summed over the vector just like we always do.

Clearly this has two indices μ and ν so we are going to give this guy μ and ν like this and now it is:

$$\nabla_{\vec{e}_\mu}\vec{e}_\nu=\Gamma^{\alpha}_{\nu\mu}\vec{e}_\alpha \tag{73}$$

This is a vector, all the indices are accounted for, it's got μ and ν and the α so for example if I was interested in $\mu=0$ and $\nu=2$ then this value would be:

$$\nabla_{\vec{e}_0}\vec{e}_2=\Gamma^{\alpha}_{20}\vec{e}_\alpha=\Gamma^{0}_{20}\vec{e}_0+\Gamma^{1}_{20}\vec{e}_1+\Gamma^{2}_{20}\vec{e}_2+\Gamma^{3}_{20}\vec{e}_3 \tag{74}$$

That is a vector, it's got all the components it needs. That's our definition that allows us to bridge between the CFREE notation and the comp notation and then we could push this a little bit further and we could, how would we go all the way I guess we should go out we should just go halfway so let me redo this:

$$\nabla_{B^\mu\vec{e}_\mu}A^\nu\vec{e}_\nu=B^\mu A^\nu{}_{,\mu}\vec{e}_\nu+A^\nu B^\mu\Gamma^{\alpha}_{\nu\mu}\vec{e}_\alpha \tag{75}$$

There's a lot of dummy indices here so I want to pull out this vector here $\vec{e}_\alpha$ and I want to pull out the B^μ as well so I need to change the $\alpha\rightarrow\nu$ and the $\nu\rightarrow\delta$:

$$\nabla_{B^\mu\vec{e}_\mu}A^\nu\vec{e}_\nu=B^\mu A^\nu{}_{,\mu}\vec{e}_\nu+A^\delta B^\mu\Gamma^{\nu}_{\delta\mu}\vec{e}_\nu \tag{76}$$

I would end up with:

$$\nabla_{B^\mu\vec{e}_\mu}A^\nu\vec{e}_\nu=\left[A^\nu{}_{,\mu}+\Gamma^{\nu}_{\delta\mu}A^\delta\right]B^\mu\vec{e}_\nu \tag{77}$$

Then in true component notation we eliminate $\vec{e}_\nu$ this altogether and this ends up just being:

$$\nabla_{B^\mu\vec{e}_\mu}A^\nu=B^\mu A^\nu{}_{\|\mu}=X^\nu \tag{78}$$

Then ν has an unpaired contravariant index because we abandon $\vec{e}_\nu$ so there's no pairing for it and we know that anything with an unpaired contravariant index is just what we call it a vector in the pure component notation. That's how we make that link out to the computation. One thing you should notice is that MTW points out is that the ordering μ,ν flips and you go with ν,μ on this definition, I'm not sure why I am sure that eventually if we work with this enough we'll see that that's a convenient choice but I'm not quite sure why that is.

With all this in mind we can quickly do the rundown of the different notations we're dealing with I can't get away from liking to do that so we have the Covariant derivative in pure comp notation $A^\nu{}_{\|\mu}$ and the Covariant derivative in the direction of $\vec{B}$ looks like $B^\mu A^\nu{}_{\|\mu}$. This guy in the abstract notation looks like that $\nabla_{\vec{B}}\vec{A}$ and the first guy in the abstract notation would just look like $\nabla\vec{A}$. That in the MTW machine notation it's $\nabla\left(\,,\vec{A}(x),\,\right)$ and $\nabla\left(\,,\vec{A}(x),\vec{B}\right)$. The last one can also be abbreviated by overloading with just $\nabla(A,B)$ and that's the same thing.

Ultimately given that this thing $\nabla_{\vec{B}} \vec{A}$ is a vector the presumption is that if you put in some 1-form here you could turn this into $\langle \omega, \nabla_{\vec{B}} \vec{A} \rangle$ which is a real number. Again this notation is very rarely used, it's more of an expository thing to sort of teach you how to get to this point $\nabla_{\vec{B}} \vec{A}$ and those are the different varieties of notation that we're going to be dealing with.

While we're here why don't we have a quick look at something that will ultimately become very important:

$$\nabla_{\vec{B}} \vec{A} - \nabla_{\vec{A}} \vec{B} \tag{79}$$

We already hinted about this in the last lecture but if we were to do this and break this down into the comp notation, if we just do the full break down into the comp notation, what would this look like?

$$\begin{aligned} \nabla_{\vec{B}} \vec{A} - \nabla_{\vec{A}} \vec{B} &= B^{\nu} \left[A^{\mu}{}_{,\nu} + \Gamma^{\mu}_{\gamma\nu} A^{\gamma} \right] \partial_{\mu} - A^{\nu} \left[B^{\mu}{}_{,\nu} + \Gamma^{\mu}_{\gamma\nu} B^{\gamma} \right] \partial_{\mu} \\ &= \left[B^{\nu} A^{\mu}{}_{,\nu} - A^{\nu} B^{\mu}{}_{,\nu} + B^{\nu} \Gamma^{\mu}_{\gamma\nu} A^{\gamma} - A^{\nu} \Gamma^{\mu}_{\gamma\nu} B^{\gamma} \right] \partial_{\mu} \end{aligned} \tag{80}$$

Then let's see what happens next. Now we look at these two last terms and we're going to make it so that they're indices are comparable and the way we do that is we are going to exchange the indices, we want to make the indices of the B the same in both of these two last terms which means I have to sort of make a change. Now the good news is that the μ and the γ are dummy indices so I can name them anything I want so I'm going to change $\nu \leftrightarrow \gamma$. You understand, this is a dummy index I can literally name it anything I want so there's really no trouble doing this so this is going to end up being:

$$\nabla_{\vec{B}} \vec{A} - \nabla_{\vec{A}} \vec{B} = \left[B^{\nu} A^{\mu}{}_{,\nu} - A^{\nu} B^{\mu}{}_{,\nu} + B^{\nu} \Gamma^{\mu}_{\gamma\nu} A^{\gamma} - A^{\gamma} \Gamma^{\mu}_{\nu\gamma} B^{\nu} \right] \partial_{\mu} \tag{81}$$

That is exactly the same as (80), there's simply just renaming indices but what's nice is I can now kind of compare these two things and I end up with, well everywhere now B got ν on these last two terms. The question is now is there anything further we can do to simplify this last two term? The last two terms will cancel if $\Gamma^{\mu}_{\gamma\nu} = \Gamma^{\mu}_{\nu\gamma}$, in other words if the connection coefficient is symmetric in the lower two indices, if that were true then the last two terms would cancel and all that would be left is the first two terms. It turns out that it will become another assumption of General Relativity that this connection that all connections involved in General Relativity have this property that the connection is symmetric in those lower two indices and this property is called Torsion free that means the Torsion tensor equals zero. We haven't defined the Torsion tensor but eventually we will it's just too tempting to go this far and not at least demonstrate this little point. It's also a bit of foreshadowing that will help make later stuff a little unsurprising so ultimately we want this property we want all General Relativity to have torsion free connections. There are alternative theories to General Relativity that do have torsion full connections, we're not going to discuss any of those and there's really no particular evidence that there are very useful theories, that is, there's nothing we know of that isn't pretty well accounted for by regular General Relativity. That's an arguable statement, I'm sure, but this is a class on standard General Relativity which always involves these torsion free environments which is an assumption of the theory ultimately. This thing here ultimately equals:

$$\nabla_{\vec{B}} \vec{A} - \nabla_{\vec{A}} \vec{B} = \left[B^{\nu} A^{\mu}{}_{,\nu} - A^{\nu} B^{\mu}{}_{,\nu} \right] \partial_{\mu} = [B, A] \tag{82}$$

Notice that these (right hand side) do not have Covariant derivatives, these (left hand side) are Covariant derivatives in certain directions but ultimately in a torsion free environment it's actually equal to the difference of two terms involving partial derivatives and we give that a symbol we call that $[B, A]$ the Lie Bracket of $\vec{B}$ and $\vec{A}$. We put $\vec{B}$ first because of the order of the left hand side of (82). You should know that the commutator of being would, in operator form would be:

$$[B, A]=BA-AB \qquad (83)$$

Which is opposite of the commutator $[A, B]$ so it does matter but in normal operator notation that's what you get here we're kind of mimicking that because the $\vec{B}$ is here first. That's an important little extra thing to do so that that actually basically just finishes up our analysis of of this part of the CFREE notation and notice how we always really when we develop this notation, the CFREE free notation, we kind of always go back and lean on the comp stuff to help us understand it and that I think is pretty much a universal facet of General Relativity CFREE. You can do a lot of interesting things in CFREE notation and it's very very useful to speak of things in generality that way but when he gets down to the grind you always kind of go back to comp notation and you really need to be able to look at CFREE notation and immediately convert it to comp notation in your head, well not literally in your head.

The comp notation seems to be intimidating but we actually work hard to understand it, you can leverage all that hard work you did to understand the comp notation and you should be able to immediately apply it to the abstract notation without having to sort of recreate an entire new analysis of the subject just in the abstract although we are going to do that, we're always going to be making this conversion leaning on this conversion to understand things.

What will we talk about next time and so the next topic that we are going to get into, not in this lesson but the next topic is going to the topic of parallel transport using the CFREE notation so the CFREE parallel transport, that's next and the basic principle in CFREE parallel transport is basically that the Covariant derivative of a Vector field $\vec{A}$ along some curve $\gamma(\tau)$ whose tangent vector is $\dot{\gamma}(\tau)$, This is what we've been calling $\vec{B}$ here, it's very simple it's just the tangent vector of a curve. Whenever we put in $\vec{B}$ we were implicitly assuming that $\vec{B}$ was the tangent vector for some curve. Now we're gonna instead of talking mostly about the vector $\vec{B}$, we're gonna emphasize the curve and we're gonna specify the tangent vector of a curve so now we're going to say $\nabla_{\dot{\gamma}(\tau)}\vec{A}$ so we have some point P there's a neighborhood of P , on this neighborhood of P we have a Vector field which is $\vec{A}$ and $\vec{A}$ has a value at P and it has a value along this curve as well, It may have changed a lot or it may have change a little bit, the question is how does the Vector field change along this curve γ ? The key to that change is not the whole curve γ itself but just the tangent vector of that curve at every spot and that is what we mean by this term $\nabla_{\dot{\gamma}(\tau)}\vec{A}$ so we're emphasizing the curve and when that Covariant derivative of $\vec{A}$ with respect to this curve equals zero then we say that the Vector field $\vec{A}$ is auto parallel on γ and so we're gonna start talking about that in our next lesson and re-understand the notion of parallel transport in terms of these abstract coordinate free notation so see you next time.