Lecture 3

In this lecture we want to talk a little more about index notation and specifically 4 vectors as well as tangent vectors dual vectors and tensors? When vectors are timelike space like and null like.

9 4-Vectors

In this lecture we will be introducing different vectors in the context of curved spaces/spacetimes and seeing how they transform. When we talk of vectors it is most useful to talk about the **components** of vectors given we have picked some basis. While this is the most convenient to out human brains, it is not how the theory of GR describes reality. In GR, as I have hopefully stressed so far, everything is fundamentally co-ordinate independent. So when we talk of vectors we are really talking about the objects V which can be expanded in terms of a basis b_μ with components V^μ:

$$V := V^\mu b_\mu \tag{5}$$

Note while we are using index notation for the two objects here it is not the same for both of them. For each $\mu = 0, 1, ..., d-1$, the component V^μ is just a number, where as each b_μ is a basis vector with components $\left(b_\mu\right)^i = \delta^i_\mu$. We mainly will be dealing with the quantity V^μ but it is always implied implicitly that the real vector is given by 5 above.

Recall in mechanics when we want to discuss vectors we take a derivative with respect to time of some co-ordinates and this gives us the perfect definition of a vector as a arrow at a point pointing in the direction of motion. In a similar manner we can take our 4-vector and take a derivative. Though before our derivatives were with repsect to time, now that we have relativity time is no longer a background parameter. We will need to is to consider the trajectory that the particle is on. This is given by $x^\mu(\lambda)$ is a map from $\mathbb{R} \to \mathcal{M}$ that paramatrizes the trajectory, λ here is just any parameter that paramatrizes the curve. For example for massive particles we can always choose their proper time $\tau = x^2$, that is the time as seen in their reference frame, to paramatrize their trajectories. With this formalism of describing trajectories we can write down the definition of a velocity 4 vector cleanly as:

$$V^\mu := \frac{dx^\mu}{d\lambda} \tag{6}$$

This definition might seem a little strange as it is very dependant on the choice of parametrization, which seems arbitrary (for example we can scale lambda by any factor and get the same curve if we mess around with the end points). To pick out a paramatrization, a normalization condition such that the vector has length one, (where by length I mean inner product with respect to the metric) can be applied to the vector. This condition is equivalent to paramatrizing by proper time, (or proper distance if the curve is space-like) which is conveniant.

With this definition we can consider what happens if we change co-ordiante system. We had been using the co-ordinate system $x^\mu(\lambda)$ but we can always choose a different co-ordinate system $y\mu(\lambda)$ and this should not change the physics. When we make a change of co-ords we always have it that the two different co-ord are functions of each other, (Think of $x = r\cos\theta$). Hence we can apply the chain rule to the definition 2. Partial derivatives do not obey the chain rule, they instead have a sum of "chain rule like terms", but this is exactly index notation!, so in essence the chain rule trick of "cancling the "dx's" is restored with index notation:

$$\frac{dx^\mu(\lambda)}{d\lambda} = \frac{dy^0(\lambda)}{d\lambda}\frac{\partial x^\mu(\lambda)}{\partial y^0(\lambda)} + \cdots + \frac{dy^{d-1}(\lambda)}{d\lambda}\frac{\partial x^\mu(\lambda)}{\partial y^{d-1}(\lambda)} = \frac{dy^\mu(\lambda)}{d\lambda}\frac{\partial x^\mu(\lambda)}{\partial y^\mu(\lambda)} \tag{7}$$

We can then notice that $\frac{dy^\nu}{d\lambda}$ is exactly the same as the definition of V_ν in 6, just in the new coordinate system. This means we can call this vector $(V')^\nu$ and relate to V^μ by manipulating the above expression 7:

$$(V')^\nu = \frac{\partial y^\nu}{\partial x^\mu} V^\mu \tag{8}$$

This defines how these type of vectors transform, we will call such vectors **contravariant**. It is possible to conceive of another type of vector. If we had a scalar function $f(x)$ on our manifold, that is some function $f : \mathcal{M} \to \mathbb{R}$, then we can get a vector from this quantity by taking derivatives of this function with respect to all the components:

$$W_\mu = \frac{\partial f(x)}{\partial x^\mu}$$

This defines a different type a vector as if we consider the same transformation as before, the chain rule now tells us:

$$\frac{\partial f(x)}{\partial x^\mu} = \frac{\partial f(x)}{\partial y^0}\frac{\partial y^0}{\partial x^\mu} + \cdots + \frac{\partial f(x)}{\partial y^{d-1}}\frac{\partial y^{d-1}}{\partial x^\mu} = \frac{\partial f(x)}{\partial y^\nu}\frac{\partial y^\nu}{\partial x^\mu}$$

Which means if we call this transformed vector $(W')_\nu$, then this is related to W_μ by:

$$(W')_\nu = \frac{\partial x^\mu}{\partial y^\nu} W_\mu \tag{9}$$

Note the key difference between this and 9 is that the positions of x and y in the derivatives have swapped. We call any vector that transforms this way a **co-variant** vector.

We are now able to see the most important property of objects that transform like this. Let me take a contravariant vector V^μ and some other co-variant vector W_ν and contract them together with the same index (this is just like a dot product). This is just the quantity $W_\mu V^\mu$. What is special about this quantity is that when we make any co-ordinate transformation, such as $x^\mu \to y^\mu(x)$, then this object transforms via:

$$W_\mu V^\mu \to (W')_\mu (V')^\mu = \frac{\partial x^\alpha}{\partial y^\mu} W_\alpha \frac{\partial y^\mu}{\partial x^\beta} V^\beta$$

But we can then use our knowledge of the chain rule, to write:

$$(W')_\mu (V')^\mu = \frac{\partial x^\alpha}{\partial x^\beta} W_\alpha V^\beta = \delta^\alpha_\beta W_\alpha V^\beta = W_\alpha V^\alpha$$

Where we have used the fact that $\frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta$. This is a remarkable result, the quantity $W_\mu V^\mu$ is invariant under co-ordinate transformations. Since we have been so general, this tells us the works for **any** vectors that transform in the same way as these co-vectors and contra-variant vectors do. What makes this remarkable is if we recall what we discussed in Lecture 1, one of the postulates of general relativity is the principle of general co-variance, in that physics should not depend on the co-ordinates used to describe it. Einstein then concluded that in order to build a theory of gravity that satisfies this principle, he must construct it out of objects that transform this way.

10 Tensors

We are now in a position to discuss the building blocks of General Relativity, which are Tensors. With the way we have set things up these should present themselves as canonical extensions of co-vectors and contravariant vectors, despite their abstract definition.

10.1 Tensor definition:

A tensor of type $(k, 0)$ and rank k, given a co-ordinate system x^μ is defined as he object:

$$T^{\mu_1,\ldots,\mu_k}(x)$$

Providing that this object under co-ordinate transformations, $x^\mu \to y^\mu$, via:

$$T^{\mu_1,\ldots,\mu_k}(x) \to (T')^{\nu_1,\ldots,\nu_k}(y) = \frac{\partial y^{\nu_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\nu_k}}{\partial x^{\mu_k}} T^{\mu_1,\ldots,\mu_k}(x)$$

Similarly we can also define a tensor of type $(0, s)$ which must transform via:

$$T_{\mu_1,\ldots,\mu_s}(x) \to (T')_{\nu_1,\ldots,\nu_s}(y) = \frac{\partial x^{\mu_1}}{\partial y^{\nu_1}} \cdots \frac{\partial x^{\mu_s}}{\partial y^{\nu_s}} T_{\mu_1,\ldots,\mu_s}(x)$$

It is also possible to have tensors which have a mix of upper and lower indices, in this case each index transforms according to whether it is upper or lower. The rank of a matrix is the total number of indices i,e, $k+s$. It is easy to see from these definitions, that when someone asks what a Tensor is, you may simply reply "A Tensor is something that transforms like a Tensor". While being purposely obfuscating this is entirely accurate.

11 The Metric Tensor

Having discussed Tensors in quite and abstract manner we can now take the time to study the simpler case of a rank 2 tensor, and probably the most important tensor in GR. We have seen already how the metric is a matrix that tells us how to measure distances on surfaces that can possibly be curved. We can then extend this to measure the length of any vectors on possible curved space. The natural extension is to take the standard inner product and replace what would be a $\delta_{\mu\nu}$ with a $g_{\mu\nu}$:

$$V^2 = V^\mu g_{\mu\nu} V^\nu$$

For example in flat spacetime when we have the Minkowski metric, $g_{\mu\nu} = \eta_{\mu\nu}$, then this is just:

$$V^2 = (V^0, V^1, V^2, V^3) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} = -(V^0) + (V^1)^2 + (V^2)^2 + (V^3)^2$$

What we can then note is that if we employ the physical constraint, that the length of this vector should be independent of the choice of co-ordinates. Then we should be able to transform the co-ordinates $x^\mu \to y^\mu$ (which means the vector V^μ transforms via 8), which leads to the following condition:

$$V^2 = V^\mu g_{\mu\nu}(x) V^\nu \to (V')^\mu g'_{\mu\nu}(y) (V')^\nu = \frac{\partial y^\nu}{\partial x^\alpha} V^\alpha g'_{\mu\nu}(y) \frac{\partial y^\mu}{\partial x^\beta} V^\beta = \frac{\partial y^\nu}{\partial x^\alpha} V^\alpha g'_{\mu\nu}(y) \frac{\partial y^\mu}{\partial x^\beta} V^\beta$$

So if $(V')^2 = V^2$, the condition on the metric $g'_{\mu\nu}$ is that:

$$g_{\mu\nu}(x) = \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} g'_{\alpha\beta}(y)$$

Or equivalently:

$$g_{\alpha\beta} \to g'_{\alpha\beta}(y) = \frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\beta} g_{\mu\nu}(x) \tag{10}$$

Meaning the metric must be a rank 2-tensor! We can also notice, that if we take just the two terms: $V^\mu g_{\mu\nu}$, then this quantity transforms in the following way:

$$V^\mu g_{\mu\nu} \to (V')^\mu g'_{\mu\nu} = \frac{\partial y^\mu}{\partial x^\gamma} V^\gamma g_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} = \delta^\alpha_\gamma V^\gamma g_{\alpha\beta} \frac{\partial x^\beta}{\partial y^\nu} = \frac{\partial x^\beta}{\partial y^\nu} V^\alpha g_{\alpha\beta}$$

I.e. this quantity transforms like a co-vector with one lower index. Precisely for this reason this quantity is denoted

$$V_\nu := g_{\nu\mu} V^\mu \tag{11}$$

We can notice one more fact about the metric tensor. If we combine the fact that it must be symmetric, with the fact that it must have non zero eigenvalues[1], then we can conclude that the metric tensor g must have an inverse g^{-1}. This inverse can be denoted $g^{\mu\nu}$ and then we have that:

[1] This follows from the fact that the only vector with zero length is the zero vector.

$$g_{\mu\nu}g^{\nu\gamma} = \delta_{\mu}^{\gamma}$$

This allows us to manipulate expressions by acting with the metric or its inverse to expressions with index's. For example if we act on both sides of 11 with $g^{\gamma\nu}$:

$$g^{\gamma\nu}V_{\nu} = g^{\gamma\nu}g_{\nu\mu}V^{\mu} = \delta_{\mu}^{\gamma}V^{\mu} = V^{\gamma}$$

In this sense we think about the metric as an object that raises or lowers indices.