May 12th Meeting Notes

May 11, 2019

Spacetime, coordinate systems, and tensors

- General relativity is based on the idea that the world of four-dimensional spacetime is curved rather than Euclidean.
- The curvature (of both space and time) is determined by the distribution of mass/energy.
- This means that the curvature is potentially different at every point in spacetime.
- So at each point in spacetime we may have a *different coordinate system.*
- Because of all this, we need to write equations in such a way that they will be valid in any coordinates.
- Such equations are called *covariant* and the idea behind this is referred to as *general covariance.*
- This is the whole reason we use tensors. Tensors "transform properly" and "mean the same thing" in different coordinate systems (or frames). Therefore, equations written entirely in tensors are automatically covariant. They are valid at all points in spacetime.

The lack of Cartesian reference coordinates

- In previous work (Mechanics and E&M, for example) we had coordinate systems like (2D) polar and (3D) spherical, etc. In all of these cases we could use geometric arguments to establish a relationship to Cartesian coordinates.
- So when things got complicated (like basis vectors changing) we could fall back to Cartesian. We were never forced to work *completely* in any of the other systems.
- We don't have this luxury in GR for two reasons:
  1. The translations might be very hard to work out for arbitrary curvature. Plus, we'd potentially need to do it individually for every point in spacetime.
  2. We are dealing with *intrinsic* curvature and (apparently) some intrinsically curved spaces can't be embedded in a higher dimensional Euclidean space.

Vectors and tangent spaces

- At each individual point in spacetime we can imagine a *tangent space.* The tangent space is flat rather than curved.
- You *can* draw an infinitesimal vector (or differential displacement) in curved spacetime because it has no length, only direction. You *can't* draw a finite-length vector (or displacement) in curved spacetime, because it would "stick out" of the curved space.
- However, you *can* draw a finite vector in the tangent space that intersects the curved space at the tail of the vector.
- So when dealing with vectors you have the choice of thinking of them as finite and in the tangent space or as differential and in curved spacetime.

Some tensor basics

Remember that, officially: $[x^\mu]$ is the upper index vector x, whereas x^μ is the μth component of that vector. We don't always adhere to that distinction. But I'll try to remember to do it in what follows.

$[v^\mu]$ is called a *vector.*

It might be the tangent vector along a parameterized path *in* spacetime.

$[\phi_\mu]$ is called a *one-form.*

It might be the change in a scalar function *on* spacetime as we move along the path above. In other words, a gradient.

(By the way, there is no special significance to the names v and ϕ. I'm just using variable names that are evocative of the concepts in play.)

Now, assume that we have two different coordinate systems. I'll use a bar rather than a prime to distinguish them (because the prime tends to get in the way of the upper index).

In the full spacetime we have four dimensions, so in each coordinate system a vector can be indexed by μ where μ runs from 0 to 3. But we'll also look at simpler examples with fewer dimensions.

We have some transformation T which takes a vector in the unbarred system and outputs a vector in the barred system:

$$\bar{x} = T(x)$$

Bear in mind that each component in one basis will generally depend on *all* the components in the other basis:

$$\begin{aligned}
\bar{x}^0 &= T_0\left(x^0, x^1, x^2, x^3\right) \\
\bar{x}^1 &= T_1\left(x^0, x^1, x^2, x^3\right) \\
\bar{x}^2 &= T_2\left(x^0, x^1, x^2, x^3\right) \\
\bar{x}^3 &= T_3\left(x^0, x^1, x^2, x^3\right)
\end{aligned}$$

So given that we know all these transformations, we will know how a given barred object changes when the corresponding unbarred change happens. That is to say, we will know the partial derivative: $\partial \bar{x}^\mu / \partial x^\mu$.

Now we are in a position to state tensor transformations between the unbarred system and the barred system:

$$(1) \quad \bar{v}^\mu = \frac{\partial \bar{x}^\mu}{\partial x^\nu} v^\nu$$

$$(2) \quad \bar{\phi}_\mu = \frac{\partial x^\nu}{\partial \bar{x}^\mu} \phi_\nu$$

Why do we even need these if we already have the transformation T? I suppose that equations involving the partial derivatives *might* be useful, even if we don't know what their values are. But I think the main point is that these tensor transformations are what we are going to build on in order to pursue general covariance. To be able to write all our equations using only tensors.

Just to be clear: An object is a legitimate upper-indexed (rank 1) tensor if it transforms according to rule (1) above. It's a legitimate lower-indexed tensor if it transforms according to rule (2). Now we'll investigate some specific objects to see if they follow those rules.

Tensors with upper indices

Let us stipulate that: $[x^\mu]$ is a *position vector* for a parametric path in spacetime.

Its components are: $x^0(\lambda), x^1(\lambda), ...$, where λ is the scalar parameter of the path.

So the component $x^\mu = x^\mu(\lambda)$.

A differential displacement vector

Consider the differential (infinitesimal) displacement along the path $[x^\mu]$. Call a component of the displacement dx^μ. What happens when we transform this to $d\bar{x}^\mu$?

Since each barred component is a function of all the unbarred components, we have:

$$d\bar{x}^\mu = \frac{\partial \bar{x}^\mu}{\partial x^0} dx^0 + \frac{\partial \bar{x}^\mu}{\partial x^1} dx^1, ... \text{ So that } d\bar{x}^\mu = \frac{\partial \bar{x}^\mu}{\partial x^\nu} dx^\nu.$$

So the differential displacement vector is an upper-indexed tensor.

A tangent vector

Suppose that $[v^\mu]$ is the tangent vector to the position vector $[x^\mu]$ above.

Then it will have the components: $v^0 = \dfrac{dx^0}{d\lambda},\ v^1 = \dfrac{dx^1}{d\lambda}, ...$ In other words $v^\mu = \dfrac{dx^\mu}{d\lambda}$.

When we transform it, each component of $[\bar{v}^\mu]$ must end up being: $\bar{v}^\mu = \dfrac{d\bar{x}^\mu}{d\lambda}$.

However, $\dfrac{d\bar{x}^\mu}{d\lambda} = \dfrac{\partial \bar{x}^\mu}{\partial x^0}\dfrac{dx^0}{d\lambda} + \dfrac{\partial \bar{x}^\mu}{\partial x^1}\dfrac{dx^1}{d\lambda}, ... = \dfrac{\partial \bar{x}^\mu}{\partial x^\nu}\dfrac{dx^\nu}{d\lambda}$

And since $\dfrac{dx^\nu}{d\lambda} = v^\nu$, we have: $\bar{v}^\mu = \dfrac{\partial \bar{x}^\mu}{\partial x^\nu} v^\nu$.

So a tangent vector is an upper-indexed tensor.

Spacetime vs. the tangent space

With regard to the two tensors above, the tangent vector $[v^\mu]$ has a finite extent. It "points out" of curved spacetime. Therefore, it must exist in the tangent space.

The differential displacement vector $[dx^\mu]$ is also tangent to the curve $[x^\mu]$. But being infinitesimal, it can exist *in* curved spacetime.

Tensors with lower indices

Assume that ϕ is a *scalar function* on spacetime.

So it takes on a specific value at any given point, and is a function of the coordinates.

The function $\phi = \phi\left(x^0, x^1, ...\right)$.

As an example, ϕ might be the temperature along the path x^μ defined above.

Now define the $[\phi_\mu]$ to be the change in ϕ as we move along the path.

The component ϕ_μ will be: $\dfrac{\partial \phi}{\partial x^\mu}$.

When we transform this component (which depends on all of the unbarred variables) applying the chain rule gives:

$$\bar{\phi}_\mu = \frac{\partial \phi}{\partial \bar{x}^\mu} = \frac{\partial \phi}{\partial x^\nu}\frac{\partial x^\nu}{\partial \bar{x}^\mu} = \frac{\partial x^\nu}{\partial \bar{x}^\mu}\frac{\partial \phi}{\partial x^\nu} = \frac{\partial x^\nu}{\partial \bar{x}^\mu}\phi_\nu$$

So the gradient of a scalar function is a one-form.

Higher order tensors

Higher order tensors can be viewed as the *direct product* of vectors and one-forms. For example:

$$Z^{\alpha\beta} = X^\alpha Y^\beta$$

$$Z_{\alpha\beta} = X_\alpha Y_\beta$$

$$Z^\gamma{}_{\alpha\beta} = W^\gamma X_\alpha Y_\beta$$

All we are doing here is directly multiplying individual components to form the higher-rank tensor. For those who recall the outer product that we used in quantum information – this is similar, but with no conjugation.

So the transformations for the higher-rank tensors follow directly from the transformations of their "factors."

$$\bar{Z}^{\alpha\beta} = \bar{X}^\alpha \bar{Y}^\beta = \frac{\partial \bar{x}^\alpha}{\partial x^\rho}X^\rho \frac{\partial \bar{x}^\beta}{\partial x^\sigma}Y^\sigma = \frac{\partial \bar{x}^\alpha}{\partial x^\rho}\frac{\partial \bar{x}^\beta}{\partial x^\sigma}Z^{\rho\sigma}$$

$$\bar{Z}_{\alpha\beta} = \bar{X}_\alpha \bar{Y}_\beta = \frac{\partial x^\rho}{\partial \bar{x}^\alpha}X_\rho \frac{\partial x^\sigma}{\partial \bar{x}^\beta}Y_\sigma = \frac{\partial x^\rho}{\partial \bar{x}^\alpha}\frac{\partial x^\sigma}{\partial \bar{x}^\beta}Z_{\rho\sigma}$$

$$\bar{Z}^\gamma{}_{\alpha\beta} = \bar{W}^\gamma \bar{X}_\alpha \bar{Y}_\beta = \frac{\partial \bar{x}^\gamma}{\partial x^\tau}W^\tau \frac{\partial x^\rho}{\partial \bar{x}^\alpha}X_\rho \frac{\partial x^\sigma}{\partial \bar{x}^\beta}Y_\sigma = \frac{\partial \bar{x}^\gamma}{\partial x^\tau}\frac{\partial x^\rho}{\partial \bar{x}^\alpha}\frac{\partial x^\sigma}{\partial \bar{x}^\beta}Z^\tau{}_{\rho\sigma}$$