GR lecture 4-1

Coordinate basis, general metric

I. SOLVING EXERCISES FROM LAST WEEK

II. GENERAL CURVED COORDINATES AND METRIC

Let us now formalize some of our previous discussion of curved coordinates. We can use coordinates $x^\mu = (x^0, x^1, x^2, x^3)$ for Minkowski space other than the inertial (t, x, y, z). When these are general linear functions of (t, x, y, z), we get slanted coordinates. When these are general non-linear functions of (t, x, y, z), we get curved coordinates. In slanted coordinates, the metric $g_{\mu\nu}$ becomes a general symmetric matrix. When the coordinates are curved, $g_{\mu\nu}$ can be different at different points – it becomes a field in spacetime $g_{\mu\nu}(x)$. In Einstein's formulation of GR, this is the gravitational field. The shift from merely curved coordinates into curved spacetime consists in simply allowing $g_{\mu\nu}(x)$ to be general, rather than restricting it to be a coordinate transformation of $\eta_{\mu\nu}$.

In curved coordinates, we can no longer think of vectors as stretching between two distant points. In particular, the coordinates x^μ themselves, or coordinate differences $x^\mu - \tilde{x}^\mu$ between two points, are not vectors. The reason is that the coordinate axes bend and change across spacetime, so a vector's components between two different points do not have a consistent meaning. Every vector or tensor must live at some single spacetime point! However, we can still talk of infinitesimal displacement vectors dx^μ, as well as about covectors $\partial f/\partial x^\mu$ constructed as the gradients of scalar functions at the point x. Equivalently, we can think of an infinitesimal vector $\varepsilon^\mu = (\varepsilon^0, \varepsilon^1, \varepsilon^2, \varepsilon^3)$ as an arrow connecting the two nearby points (x^0, x^1, x^2, x^3) and $(x^0 + \varepsilon^0, x^1 + \varepsilon^1, x^2 + \varepsilon^2, x^3 + \varepsilon^3)$. This construction defines a particular basis of vectors at each spacetime point, which is derived from our coordinate system. We call this a coordinate basis. This may also be a good time to recall the "dual" point of view, in which dx^μ is not a vector but a basis of covectors, and $\partial/\partial x^\mu$ is not a covector but a basis of vectors. This is just another (in fact, more concise) way to express the concept of a coordinate basis.

During most of the course, we will exclusively use coordinate bases. However, in principle, the task of labelling points of spacetime with coordinates can be separated from the task of choosing a basis of vectors at each point. In fact, there exists a more powerful formulation of GR, which we might learn about later, that uses non-coordinate bases.

Now, how should we imagine a vector that sits at the point x, but it not infinitesimal? For some reason, my personal intuition never recognizes this as a problem, but yours might. It might then be helpful to imagine the curved spacetime as a curved 4d surface within a larger flat space, and to imagine the space of vectors at x as a 4d flat spacetime that is tangent to "true" curved one. Vectors along this flat "tangent space", with their "arrow's tail" at x, can be made as long or as short as we wish. However, unless they're infinitesimal, the "head" of their arrow won't lie anywhere in the true, curved spacetime. Due to this geometric intuition, the space of vectors at a point is in fact often called the "tangent space". Similarly, the space of covectors is called the "cotangent space".

Given a coordinate system x^μ, we can always transform to a different one x'^μ, again given by arbitrary non-linear functions of x^μ. Under such a transformation, not only do the coordinate labels on a spacetime point change, but so does the coordinate basis which we use for writing the components of vectors and tensors. It's easy to see that the appropriate basis transformation matrix is given by the matrix of derivatives:

$$M_\mu{}^\nu = \frac{\partial x^\nu}{\partial x'^\mu} \; ; \quad (M^{-1})_\mu{}^\nu = \frac{\partial x'^\mu}{\partial x^\nu} \; , \tag{1}$$

where we note an important subtlety about the notation: the partial derivative $\partial/\partial x^1$ is taken at fixed (x^0, x^2, x^3), while the partial derivative $\partial/\partial x'^1$ is taken at fixed (x'^0, x'^2, x'^3), which is not the same! To prove that (1) is indeed the basis transformation matrix, we need simply to write the chain rules for partial derivatives:

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu \; ; \quad \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} \; . \tag{2}$$

This proves that the components of the vector dx^μ and the covector ∂_μ indeed transform in accordance with the basis transformation matrix (1). We can now write the transformation rule for the components of an arbitrary tensor, such as:

$$T'_{\mu\nu\rho}{}^\sigma = \frac{\partial x^\kappa}{\partial x'^\mu} \frac{\partial x^\lambda}{\partial x'^\nu} \frac{\partial x^\alpha}{\partial x'^\rho} \frac{\partial x'^\sigma}{\partial x^\beta} T_{\kappa\lambda\alpha}{}^\beta \; . \tag{3}$$

In particular, the metric transforms as:

$$g'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma} \; . \tag{4}$$

Note that this rule is actually equivalent to the trick we've been using until now to transform the metric between coordinate system, e.g. in our discussion of polar and spherical coordinates.

Armed with these tools, let us consider a final example of curved coordinates in flat spacetime, which is arguably more realistic than (t, x, y, z) itself. Let's construct a simplified version of GPS! Let there be four "satellites", hovering at rest at four corners of a tetrahedron centered at the origin:

$$\mathbf{x_1} = (-1,-1,-1) \; ; \quad \mathbf{x_2} = (-1,1,1) \; ; \quad \mathbf{x_3} = (1,-1,1) \; ; \quad \mathbf{x_4} = (1,1,-1) \; . \tag{5}$$

On each satellite there is a clock, which measures the time t in the satellites' rest frame. Each satellite continually broadcasts the reading of its clock in all directions, using signals that travel at the speed of light. Thus, an observer at any point in spacetime receives four clock readings $(\tau^1, \tau^2, \tau^3, \tau^4)$ from the two satellites. These can be used as coordinates for the spacetime point!

Exercise 1.

1. *Express* $(\tau^1, \tau^2, \tau^3, \tau^4)$ *in terms of the usual* (t, x, y, z)*.*

2. *Now, for simplicity, let's restrict our attention to the symmetry axis* $(t, x, y, z) = (t, 0, 0, 0)$*. For points on this axis, find the components of the metric* $g_{\mu\nu}$ *and its inverse* $g^{\mu\nu}$ *in the* $(\tau^1, \tau^2, \tau^3, \tau^4)$ *basis. Which components vanish? Why?*

III. COUNTING DEGREES OF FREEDOM

Let's now return to the question of general metrics $g_{\mu\nu}(x)$ (curved spacetime) vs. transformations of $\eta_{\mu\nu}$ (curved coordinates in flat spacetime). First, let's consider a simpler analogous question from electromagnetism. A vanishing electromagnetic field can be described by $A_\mu = 0$, but also, through gauge symmetry, by $A_\mu = \partial_\mu\theta$. This defines a family of gauge potentials parameterized by 1 scalar function $\theta(x^\mu)$ of the 4 coordinates. However, a general gauge potential $A_\mu(x^\nu)$ is given by $\underline{4}$ functions of the coordinates – the 4 components of A_μ. Therefore, a general gauge potential $A_\mu(x)$ is <u>not</u> a gauge-transformed version of $A_\mu = 0$, and does describe a non-trivial electromagnetic field.

Now, consider a local version of this question – let us focus on A_μ and its derivatives at a given point x, and similarly for the gauge transformation parameter θ. We can expand θ around x in a Taylor series of derivatives $\theta, \partial_\mu\theta, \partial_\mu\partial_\nu\theta, \ldots$, which can all be chosen arbitrarily. From the gauge transformation $A_\mu \to A_\mu - \partial_\mu\theta$, we see that $\partial_\mu\theta$ will modify A_μ at x, $\partial_\mu\partial_\nu\theta$ will modify $\partial_\mu A_\nu$ at x, and so on. Now, let's count degrees of freedom order by order. $\partial_\mu\theta$ at x is an arbitrary vector with 4 components, which can always be chosen to cancel the 4 components of A_μ. Thus, the value of A_μ at a single point is meaningless, and can always be transformed to zero. However, at the next order, we notice that $\partial_\mu\partial_\nu\theta$ has only $4 \times 5/2 = 10$ independent components, because it's a symmetric matrix! These 10 components are not enough to cancel the $4 \times 4 = 16$ components of $\partial_\mu A_\nu$! We conclude that the "signature" of a non-trivial electromagnetic field should appear at first order in derivatives of A_μ, and that the "physical" part of $\partial_\mu A_\nu$ has $16 - 10 = 6$ components that cannot be transformed away. Of course, we recognize these as the $4 \times 3/2$ independent components of $F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]}$.

This is a good time to give some general formulas for component counting. The number of independent components of a rank-k totally antisymmetric tensor $T_{[\mu_1\ldots\mu_k]}$ in n dimensions is clearly:

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!} \ . \tag{6}$$

Somewhat less obviously, the number of components of a rank-k totally symmetric tensor $T_{(\mu_1\ldots\mu_k)}$ is:

$$\binom{n+k-1}{k} = \frac{n(n+1)\ldots(n+k-1)}{k!} \tag{7}$$

Exercise 2. *Prove this for $k = 2, 3$. Can you see how to prove the general case?*

Now, let's turn from gauge potentials back to metrics. A general metric consists of 10 functions – the 10 independent components of the symmetric matrix $g_{\mu\nu}$ – of the 4 coordinates. On the other hand, a general coordinate transformation $x^\mu \to x'^\mu(x^\nu)$ consists of 4 functions of 4 coordinates. Therefore, a general metric $g_{\mu\nu}(x)$ is <u>not</u> a coordinate-transformed version of the flat metric $\eta_{\mu\nu}$. Now, let's see what we can say about the Taylor series of derivatives around a fixed point x. By the transformation law (4), the value of $g_{\mu\nu}$ at x will be modified by the matrix of derivatives $\partial_\mu x'^\nu = \partial x'^\nu/\partial x^\mu$ (it doesn't really matter for this purpose whether we're talking about the matrix itself or its inverse). This matrix has $4 \times 4 = 16$ components – more than enough to set $g_{\mu\nu}$ to any value we want, in particular to set $g_{\mu\nu} = \eta_{\mu\nu}$. At the next order in derivatives, we find that $\partial_\mu g_{\nu\rho}$, with $4 \times 10 = 40$ components, is modified by $\partial_\mu \partial_\nu x'^\rho$, which again has $10 \times 4 = 40$ components. Thus, we have enough degrees of freedom to transform also $\partial_\mu g_{\nu\rho}$ into whatever we want, in particular to set $\partial_\mu g_{\nu\rho} = 0$. At the next order, we find that $\partial_\mu \partial_\nu g_{\rho\sigma}$, which has $10 \times 10 = 100$ components, is modified by $\partial_\mu \partial_\nu \partial_\rho x'^\sigma$, which has only $20 \times 4 = 80$ components (note that by (7), the number of components in the totally symmetric rank-3 object $\partial_\mu \partial_\nu \partial_\rho$ is $4 \times 5 \times 6/3! = 20$). We conclude that a general metric can be trivialized up to first order in derivatives, but at the second order $\partial_\mu \partial_\nu g_{\rho\sigma}$, there are $100 - 80 = 20$ components which cannot be transformed away. Thus, we expect that spacetime curvature should be described by some 20-component tensor constructed out of the metric's second derivatives. As we will learn, this is in fact the Riemann curvature tensor.

Let us compare this conclusion to our discussion of Newtonian gravity and non-inertial frames. The metric elements $g_{\mu\nu}$ are roughly analogous to velocities in the Newtonian story – in particular, we've seen that g_{ti} should be thought of as a velocity. The first derivatives $\partial_\mu g_{\nu\rho}$ are then roughly analogous to accelerations. The fact that $g_{\mu\nu}$ and $\partial_\mu g_{\nu\rho}$ can be transformed into the flat values $\eta_{\mu\nu}$ and 0 means that neither velocity nor acceleration are absolute – they can be transformed away by a choice of frame. Finally, $\partial_\mu \partial_\nu g_{\rho\sigma}$ is analogous to gradients of the acceleration, which cannot be completely transformed away, and which indicate the definite presence of a non-trivial gravitational field.