Notes on General Relativity

Aditya Vijaykumar

International Centre for Theoretical Sciences, Bengaluru, India.

E-mail: aditya.vijaykumar@icts.res.in

Contents

1   Tensors

The Newtonian equations of motion are often phrased in the language of vectors. For example, take Newton's law of gravitation which gives the gravitational force between two bodies of mass m_1 and m_2 with a separation vector $\vec{\mathbf{r}}$ pointing from m_1 to m_2

$$\vec{\mathbf{F}}_{12} = -G \frac{m_1 m_2}{|\vec{\mathbf{r}}|^3} \vec{\mathbf{r}} \quad . \tag{1.1}$$

Here, $\vec{\mathbf{F}}_{12}$ is the force on m_2 due to m_1. There are two ways of looking at this equation:

- A *coordinate-free way*, where we look at these objects as just vectors. The equations written in this form have the advantage that they don't demand any coordinate system to be chosen.

- A *coordinate-dependent way*, where we look at these component-wise.

Before diving straight in to general relativity and the Einstein equations, one should be comfortable with the basic mathematical language that the theory is written in – tensors. One can think of tensors to be a generalization of vectors; in fact, in the formalism we develop, it would be easy to see that vectos are tensors of a specific kind. Tensor equations can also be looked at in a coordinate-free and coordinate-dependent way; but it is the latter that we will choose to work with here.

1.1 Tensor Algebra

1.1.1 Manifolds

A tensor is a geometrical object defined on a *manifold*. One can define a manifold in a detailed geometric way, but we (hopefully) won't need that for the purposes of these notes. Instead, we will just say that a manifold in n-dimensions is something that *locally* looks like n-dimensional Euclidean (flat) space $\mathbb{R}^n$. As a simple illustration, let's consider a 2-sphere S^2. S^2 and $\mathbb{R}^2$ are definitely not the same globally. But a small bit of S^2 does look like $\mathbb{R}^2$. For example, even though the earth is (approximately) spherical, we as *local* humans perceieve it as being flat.

A manifold is then simply a set of points such that each point possesses a set of n-coordinates $(x^1, x^2, \ldots, x^n)$ which are all real numbers. These could be thought of as the counterpart of distances and angles in the Euclidean plane. Too add to this, it might even be impossible to cover a manifold in full by a non-degenerate coordinate system. Take a Euclidean example: plane polar coordinates have a degeneracy at the origin where ϕ becomes indeterminate, though we could make do with Cartesian coordinates in this case. But in other cases, this freedom might not be available. For example, it can be shown that there is no coordinate system that covers the whole of S^2 without degeneracy. For example, the "flat" map of the world (the so called Mercator projection) misses the poles.

As an illustrative example of the above, let's consider the stereographic projection which maps a sphere onto a plane. Let $A = (x^1, x^2, x^3)$ be the coordinates of a point on the sphere. Consider also the plane $\mathcal{P}$ defined by $x^3 = -1$. The stereographic projection involves drawing a line passing through the north pole and A and finding the intercept of this line on $\mathcal{P}$; let's denote the intercept by $A' = (y^1, y^2)$. We can show that,

$$(y^1, y^2) = \left(\frac{2x^1}{1 - x^3}, \frac{2x^2}{1 - x^3} \right) \tag{1.2}$$

As can be seen from the above equation, this mapping fails to capture the north pole itself (it maps a finite point onto ∞). For us to include the north pole, we have to redefine the steoreographic mapping by considering a line passing through south pole and A. Think about this a bit more.

Often, it is convenient to work with only one of these coordinate systems and keep track of the points that aren't included. This coordinate system is called a coordinate patch, and a set of coordinate patches that covers the entire manifold is called an atlas.

1.1.2 Curves and Surfaces

Given a manifold, we would be interested in the subset of points that form curves and surfaces. We shall often define them parametrically. For example, since a curve has one degree of freedom, it will depend only on one *parameter*; we can define it by:

$$x^a = x^a(u) \quad , \tag{1.3}$$

where u is the parameter and $x^1(u), x^2(u), \ldots, x^n(u)$ are all functions of u. Similarly, since a surface of m dimensions will have m degrees of freedom, we can define it by a set of m

parameters:

$$x^a = x^a(u^1, u^2, \dots, u^m) \quad . \tag{1.4}$$

In particular, if $m = n - 1$, the surface is called a *hypersurface*. Here, one can eliminate $n - 1$ parameters from n equations to give one equation connecting the coordinates *ie.*,

$$f(x^1, x^2, \dots, x^n) = 0 \quad . \tag{1.5}$$

Equivalently, if a point on an n-dimensional manifold is constrained to lie on a hypersurface, it should follow eq. 1.5. Similarly, points on an m-dimensional surface should follow $n - m$ such constraints.

1.1.3 Transformation of Coordinates

While writing down tensor equations, we would like to have it hold in all coordinate systems. Hence, we would like to see how one can realize coordinate transformations.

Let's consider a transformation of coordinates x^a to $x'^a = f^a(x_1, x_2, \dots, x_n)$, where f^a's are single-valued, continuous, differentiable functions. We can view this as passively assigning a point with coordinates $(x^1, x^2, \dots, x^n)$ the new coordinates $(x'^1, x'^2, \dots, x'^n)$. More succintly, we will just use the notation $x'^a = x'^a(x)$

We now want to see what the transformation will look like. Hence, we differentiate x' with respect to x – this will give us a matrix of coefficients:

$$\left[\frac{\partial x'^a}{\partial x^b} \right] = \begin{bmatrix} \frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} & \cdots & \frac{\partial x'^1}{\partial x^n} \\ \frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} & \cdots & \frac{\partial x'^2}{\partial x^n} \\ \vdots & & & \\ \frac{\partial x'^n}{\partial x^1} & \frac{\partial x'^n}{\partial x^2} & \cdots & \frac{\partial x'^n}{\partial x^n} \end{bmatrix} \tag{1.6}$$

The determinant of the above transformation matrix is called the Jacobian J' of the transformation. If we assume that J' is non-zero, we can solve eq 1.6 and find the inverse transformation $x^a = x^a(x')$. Using the product rule of determinants, the Jacobian of the inverse transformation $J = 1/J'$

How would the total differential of x'^a relate to the transformation matrix? We write,

$$\mathrm{d}x'^a = \sum_{b=1}^{n} \frac{\partial x'^a}{\partial x^b} \, \mathrm{d}x^b = \frac{\partial x'^a}{\partial x^b} \, \mathrm{d}x^b \quad , \tag{1.7}$$

where we invoke the *Einstein summation convention* to simplify our equations – we drop the summation altogether and understand that whenever an index is repeated, it should be summed over. This repeated index is also often called the *bound index* or the *dummy index*.

As an aside, we will also define the *Kronecker delta* to be a quantity which is either 0 or 1 according to

$$\delta_b^a = \begin{cases} 1 & \text{if} \quad a = b \\ 0 & \text{if} \quad a \neq b \end{cases} \implies \frac{\partial x^a}{\partial x^b} = \frac{\partial x'^a}{\partial x'^b} = \delta_b^a \quad . \tag{1.8}$$

1.1.4 Contravariant Tensors

We now want to define geometrical quantities with reference to their transformation properties under transformations described in the previous section. Before defining contravariant tensors in general, let's first define a contravariant vector.

Consider two neighbouring points on a manifold P and Q with coordinates x^a and $x^a + \mathrm{d}x^a$. This forms an infinitesimal vector $\mathbf{\vec{PQ}}$ attached at P and directed to Q, whose components in this coordinate system are $\mathrm{d}x^a$. This vector will get transformed to $\mathrm{d}x'^a$ in the new coordinate system, where $\mathrm{d}x'^a$ is given by eq. 1.7. Note that the transformation matrix in this case will be evaluated at point P.

We then define a *contravariant vector* X^a as a set of of quantities associated with point P which transform under a change of coordinates according to

$$X'^a = \frac{\partial x'^a}{\partial x^b} X^b \quad , \tag{1.9}$$

with the transformation matrix evaluated at P. This definition of the contravariant vector can then be generalized to *contravariant tensors*. A contravariant tensor of rank m in n-dimensions is a set of n^m quantities associated with a point P that follow a certain transformation law. For illustrative purposes, in case of $m = 2$, X^{ab} is a contravariant tensor if,

$$X'^{ab} = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} X^{cd} \quad . \tag{1.10}$$

The transformation properties of higher rank tensors can be obtained in an analogous manner. A special case is a tensor of rank zero – which is basically a scalar! This transforms according to $\phi' = \phi$.

1.1.5 Covariant and Mixed Tensors

Consider a real, continuous, differentiable scalar $\phi = \phi(x)$. We have seen that x^a can be written as a function of x'^a. Then, we can imagine differentiating ϕ with respect to x'^a,

$$\frac{\partial \phi}{\partial x'^a} = \frac{\partial \phi}{\partial x^b} \frac{\partial x^b}{\partial x'^a} \quad . \tag{1.11}$$

Note that this involves the inverse tranformation from a coordinate system x' to x. This motivates us to define a covariant vector X_a as a set of quantities that follows

$$X'_a = \frac{\partial x^b}{\partial x'^a} X_b \quad , \tag{1.12}$$

and analogously we can define a rank 2 covariant tensor by the law

$$X'_{ab} = \frac{\partial x^c}{\partial x'^b} \frac{\partial x^d}{\partial x'^c} X_{cd} \quad , \tag{1.13}$$

and so on for higher rank tensors. In the same spirit, we can define mixed tensors of rank 3 by the law

$$X'^a_{bc} = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \frac{\partial x^f}{\partial x'^c} X^d_{ef} \quad . \tag{1.14}$$

If a mixed tensor has contravariant rank p and covariant rank q, it is said to have *type* or *valence* (p, q).

Let's say we find two tensors X_{ab} and Y_{ab} which are equal component-by-component in a certain coordinate system *ie.* $X_{ab} = Y_{ab}$. Will the tensors still be equal in a transformed coordinate system? Yes! We can see this by multiplying on both sides by the transformation matrix. In other words, the equality of two tensors would hold true in any and all coordinate systems! We might be interested in introducing coordinate systems for the purposes of a specific problem, but tensorial equations are essentially coordinate-independent.

1.1.6 Tensor Fields

In vector analysis, a *vector field* defined in a region is the association of a vector to every point in that region. Analogously, a tensor field defined in a region is the association of a tensor with the same valence to every point in the region; *ie*,

$$P \to T^{a\cdots}_{b\cdots}(P) \quad , \tag{1.15}$$

where $T^{a\cdots}_{b\cdots}(P)$ is the value of the tensor at P. As one might expect, the tensor field is called continuous and differentiable if its coordinates are continuous and differentiable functions. The tensor field is called smooth is its components are differentiable to all orders, denoted mathematically by saying that the components are C^{∞}.

1.1.7 Elementary Operations

- $X^{a}_{bc} = Y^{a}_{bc} + Z^{a}_{bc}$

- $X^{a}_{bc} = \alpha W^{a}_{bc}$, where α is a scalar.

- A covariant tensor of rank 2, X_{ab}, is said to be symmetric if $X_{ab} = X_{ba}$. In this case, it has $\frac{n(n+1)}{2}$ independent components.

- A covariant tensor of rank 2, X_{ab}, is said to be anti-symmetric if $X_{ab} = -X_{ba}$. In this case, it has $\frac{n(n-1)}{2}$ independent components.

- Given a covariant tensor of rank 2, one can write it as a sum of a symmetric tensor and an anti-symmetric tensor,

$$X_{ab} = X_{(ab)} + X_{[ab]} \quad ,$$
$$\text{where} \quad X_{(ab)} = \frac{X_{ab} + X_{ba}}{2} \quad \text{and} \quad X_{[ab]} = \frac{X_{ab} - X_{ba}}{2} \quad .$$

- In general,

$$X_{(a_1 a_2 \ldots a_r)} = \frac{1}{r!}(\text{sum over all permutations of indices})$$
$$X_{[a_1 a_2 \ldots a_r]} = \frac{1}{r!}(\text{alternating sum over all permutations of indices}) \quad .$$

For example,

$$X_{[abc]} = \frac{1}{6!}(X_{abc} - X_{acb} + X_{cab} - X_{cba} + X_{bca} - X_{bac}) \quad . \tag{1.16}$$

- We can multiply a type (p_1, q_1) tensor and a type (p_2, q_2) tensor to get a type $(p_1 + p_2, q_1 + q_2)$ tensor, $eg.$ $X^a_{bcd} = Y^a_b Z_{cd}$

- Given a tensor of type (p, q), we can construct a type $(p-1, q-1)$ tensor by the process of contraction ie said a raised index equal to the lower index, eg $X^a_{bcd} \rightarrow_{contraction}$ $X^a_{acd} = \delta^b_a X^a_{acd} = Y_{cd}$.

1.1.8 Index-free interpretation of covariant tensor fields

In the x^a-coordinate system, let's introduce the following notation,

$$\partial_a = \frac{\partial}{\partial x^a} \quad . \tag{1.17}$$

We then define an operator X as,

$$X = X^a \partial_a \implies Xf = X^a \partial_a f = X^a (\partial_a f) \quad , \tag{1.18}$$

where f is an arbitrary real-valued function. It can be proved that the operator X is does not depend on the choice of the coordinate system and hence $X = X^a \frac{\partial}{\partial x^a} = X'^a \frac{\partial}{\partial x'^a}$. Hence, operating X on f will lead to the same result irrespective of the coordinate system used.

Since any vector at P is given by the above equation, we can think of the quantities $[\partial_a]_P$ as forming the basis for all the vectors at P, and $[X^a]_P$ being the components. The vector space of all the contravariant vectors at P is called the *tangent space* at P and it denoted by $T_P(M)$.

Given two vector fields X and Y, we can define a new vector field called the *commutator* or the *Lie bracket* of X and Y by

$$[X, Y] = XY - YX \quad . \tag{1.19}$$

How are we sure the commutator is indeed a vector field? Consider the action of the commutator on a scalar function f,

$$
\begin{aligned}
[X, Y]f &= XYf - YXf \\
&= X^a \partial_a (Y^b \partial_b f) - Y^b \partial_b (X^a \partial_a) f \\
&= \left(X^a \partial_a Y^b \partial_b - Y^b \partial_b X^a \partial_a \right) f + \left(X^a Y^b \partial_a \partial_b f - Y^b X^a \partial_b \partial_a f \right) \\
&= \left(X^b \partial_b Y^a - Y^b \partial_b X^a \right) \partial_a f \quad ,
\end{aligned}
$$

which proves that $[X, Y]$ is a vector field.

The properties of the Lie bracket are:

- $[X, X] = 0$

- $[X, Y] = -[Y, X]$

- *Jacobi Identity*: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

1.2 Tensor Calculus

1.2.1 Partial Differentiation of Tensors does not yield a Tensor

Consider,

$$\partial'_c X'^a = \frac{\partial}{\partial x'^c}\left(\frac{\partial x'^a}{\partial x^b}X^b\right) \tag{1.20}$$

$$= \frac{\partial x^d}{\partial x'^c}\frac{\partial}{\partial x^d}\left(\frac{\partial x'^a}{\partial x^b}X^b\right) \tag{1.21}$$

$$= \frac{\partial^2 x'^a}{\partial x^b \partial x^d}\frac{\partial x^d}{\partial x'^c}X^b + \frac{\partial x^d}{\partial x'^c}\frac{\partial x'^a}{\partial x^b}\frac{\partial X^b}{\partial x^d} \tag{1.22}$$

This exercise clearly shows that the partial derivative *does not* transform as a tensor.

Why is this so? Taking a partial derivative really means taking the difference between two points P and Q, dividing it by the separation, and taking the limit as separation goes to zero. The transformation law of tensors depends on each point, hence the difference of two vectors can have different magnitudes in different coordinate systems. So, the difference of two vectors *is not a vector*, and hence the derivative of a tensor does not transform as a vector in general.

1.2.2 The Lie Derivative

A *congruence of curves* is defined as a set of curves such that exactly one curve passes through a given point on the manifold. This is denoted by $x^a = x^a(u)$. Given this, we can define the tangent vector field $\frac{\mathrm{d}x^a}{\mathrm{d}u}$ along the curve. Doing this for every curve in the congruence, we end up with a vector field X^a given by $\frac{\mathrm{d}x^a}{\mathrm{d}u}$ at each point and defined over the whole manifold.

Conversely, we can think of a vector field $X^a(u)$ defined on the manifold and can then define a congruence of curves on the manifold, called *orbits* or *trajectories* of X^a. These curves can be obtained by solving the equation

$$\frac{\mathrm{d}x^a}{\mathrm{d}u} = X^a(x(u)) \quad . \tag{1.23}$$

Say we have an arbitrary tensor $T^{a\cdots}_{b\cdots}$ and quantities T_P and T_Q defined at two points P and Q.