

Tensors: An overview for Relativity

V.H. Belvadi¹

February 2019

¹ E-mail: hello@vhbelvadi.com

Download: <http://bit.ly/2RBpXVh>

Since in General Relativity we use curvilinear coordinates² and tensors heavily it is often convenient and preferred to start courses in relativity with an overview of these ideas. In this lecture we will attempt to lay a foundation of mathematical preliminaries that we shall make use of in later studies of both Special and General Relativity.

² You will see in your studies of G.R. that flat (Cartesian or rectangular) space is merely localised curved space. This is why we use curvilinear coordinates.

Curvilinear coordinates

Defining basic transformations

We know that every coordinate marking some point (x, y, z) in space may be expressed as functions of some (u_1, u_2, u_3) as follows:

$$\begin{aligned}x &\equiv x(u_1, u_2, u_3) \\y &\equiv y(u_1, u_2, u_3) \\z &\equiv z(u_1, u_2, u_3)\end{aligned}\tag{1}$$

On solving them we end up with the following:

$$\begin{aligned}u_1 &\equiv u_1(x, y, z) \\u_2 &\equiv u_2(x, y, z) \\u_3 &\equiv u_3(x, y, z)\end{aligned}\tag{2}$$

Of course we are making plenty of assumptions here. For example we assume that our functions all yield just one possible result for a given argument; we also assume that the function is continuous. These are in no way crippling our argument, however, since we are free to define our domain and range in such a way that these assumptions remain valid. Equations (1) and (2) are transformation equations that take us from the **rectangular coordinates** of a point, given by (x, y, z) , to its **curvilinear coordinates**, given by (u_1, u_2, u_3) .

Unit vectors

Since curvilinear coordinates need not be straight it begs the question as to what unit vectors in curvilinear coordinates are. Simply put they are unit-length vectors pointing in the direction of increase of the coordinate axes from a given point.

Say we have a point P defined by the position vector $r \equiv r(u_1, u_2, u_3)$. We are aware that $\frac{\partial r}{\partial u_1}$ gives the tangent at that point along u_1 (likewise for u_2 and u_3). We can then define a unit vector along this direction as simply $e_1 = \frac{\partial r}{\partial u_1} / \left| \frac{\partial r}{\partial u_1} \right|$ or, written more simply, $\frac{\partial r}{\partial u_1} = h_1 e_1$ (likewise we have e_2 for u_2 and e_3 for u_3).

We can do the same for the surfaces too. For example we can define some c_1 formed by u_2 and u_3 (and similarly c_2 and c_3) perpendicular to which points the unit vector E_1 (and similarly E_2 and E_3). Here all E_i like all e_i are mutually perpendicular.

Since we are working with curvilinear coordinates³ rather than linear/rectangular ones we need to remember that the unit vectors, whether e_i or E_i , change from point to point. While $P + \delta P$ may not have the same unit vectors as P , it will, just like P and like any other point in our curvilinear coordinate system, have two sets of unit vectors: e_i tangential to the coordinate curves, and E_i perpendicular to the coordinate surfaces.

Contravariant and covariant components

Let us write down our unit vectors in a slightly different form. Consider the equations

$$\mathbf{A} = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2 + A_3 \mathbf{e}_3 = a_1 \mathbf{E}_1 + a_2 \mathbf{E}_2 + a_3 \mathbf{E}_3$$

where A_i and a_i are different forms of the components of \mathbf{A} depending on which system⁴ we are working in.

We can write these equations in yet a different form. We know that $e_i = \frac{\partial \mathbf{r}}{\partial u_i}$ and, say, $E_i = \nabla u_i$. We call the right-hand terms in these equations **base vectors**. They are like e_i and E_i in every way except that they are not of unit length. Our above equations now become

$$\mathbf{A} = C_1 \frac{\partial \mathbf{r}}{\partial u_1} + C_2 \frac{\partial \mathbf{r}}{\partial u_2} + C_3 \frac{\partial \mathbf{r}}{\partial u_3} \quad (3)$$

$$\mathbf{A} = c_1 \nabla u_1 + c_2 \nabla u_2 + c_3 \nabla u_3 \quad (4)$$

where C_i are known as the **contravariant components** of the vector \mathbf{A} and c_i are known as its **covariant components**.

Arc length

The measure of the length of an arc is an important feat in any coordinate system. Indeed the measure of the shortest distance between two points is fundamental to studying almost anything in space and this is no different. For some $\mathbf{r} \equiv \mathbf{r}(u_1, u_2, u_3)$ in rectangular coordinates we have the straightforward relation

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = \sum_i \left(\frac{\partial \mathbf{r}}{\partial u_i} \right)^2 (du_i)^2 = \sum_i h_i^2 du_i^2$$

where we have used our previous definition $h_i e_i = \left| \frac{\partial \mathbf{r}}{\partial u_i} \right|$ but we omit the e_i from the final term because $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Of course all this is well-known as a form of the distance formula in rectangular coordinates. What we need to discuss is its form in a curvilinear system.

We begin with the understanding that the same definitions as above are valid, namely, $ds^2 = d\mathbf{r} \cdot d\mathbf{r}$ is the arc length, and that

³ What would be the relationship between e_i and E_i if we were working in a rectangular coordinate system?

⁴ It should be clear to you by now that a system here refers to a coordinate system and a coordinate system is defined by the unit vectors with which we choose to build them.

$d\mathbf{r} = \sum_i \frac{\partial \mathbf{r}}{\partial u_i} du_i$ so together these results give us

$$\begin{aligned} ds^2 &= d\mathbf{r} \cdot d\mathbf{r} \\ &= \sum_{i=1}^3 \sum_{j=1}^3 \left(\frac{\partial \mathbf{r}}{\partial u_i} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \right) (du_i \cdot du_j) \\ \implies ds^2 &= \sum_{i=1}^3 \sum_{j=1}^3 g_{ij} du_i du_j \end{aligned} \quad (5)$$

where $g_{ij} = \frac{\partial \mathbf{r}}{\partial u_i} \cdot \frac{\partial \mathbf{r}}{\partial u_j}$ are known as the **metric coefficients**. When orthogonal this simply means $g_{ij} = \delta_{ij}$. More important, there is no reason to restrict ourselves to $i = 1, 2, 3$ just because our brain can only visualise three dimensions. Mathematically, nothing prevents us from going as high as we want to but in four-dimensional space in particular, and therefore in relativity, this metric, represented with greek indices as $g_{\mu\nu}$, is of particular importance.

The volume element

Measuring a volume in space is our next step after measuring arc lengths. As with any vector⁵ volume is given in terms of three mutually perpendicular vectors \mathbf{a} , \mathbf{b} and \mathbf{c} about some point as $|\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}|$. Therefore,

$$\begin{aligned} dV &= \left| \left(\frac{\partial \mathbf{r}}{\partial u_1} du_1 \right) \cdot \left(\frac{\partial \mathbf{r}}{\partial u_2} du_2 \right) \times \left(\frac{\partial \mathbf{r}}{\partial u_3} du_3 \right) \right| \\ &= |(h_1 du_1 \mathbf{e}_1) \cdot (h_2 du_2 \mathbf{e}_2) \times (h_3 du_3 \mathbf{e}_3)| \\ \implies dV &= h_1 h_2 h_3 du_1 du_2 du_3 \end{aligned} \quad (6)$$

since $|\mathbf{u}_1 \cdot \mathbf{u}_2 \times \mathbf{u}_3| = 0$ gives the volume of an element in curvilinear space formed by the above three vectors.

Basics of curvilinear coordinate systems

We will end our non-tensorial discussion with an overview of some ideas already familiar to you. First, recall that the **spherical** coordinates (r, θ, ϕ) can be written in terms of (x, y, z) as follows:

$$x = r \sin \theta \cos \phi \quad y = r \sin \theta \sin \phi \quad z = r \cos \theta$$

where $0 \leq r$, $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$.

Simpler than this is the **cylindrical** (ρ, ϕ, z) system of coordinates:

$$x = \rho \cos \phi \quad y = \rho \sin \phi \quad z = z$$

where $0 \leq \rho$, $0 \leq \phi < 2\pi$ and $-\infty < z < \infty$. This is as useful in two dimensions as in three but we will not see too much of this either way while studying relativity.

Finally let us take a quick look at the triumvirate of vector operations gradient, divergence and curl, in a general curvilinear form. We

⁵ I say 'As with any vector. ...' since any definition you studied with respect to vectors in the past were universal—which is why we expect them to hold up here too.

are setting up the stage to expand this later using tensors. We have,

$$\begin{aligned}\nabla \varphi &= \text{grad } \varphi = \sum_i \frac{1}{h_i} \frac{\partial \varphi}{\partial u_i} e_i \\ \nabla \cdot \mathbf{A} &= \text{div } \Phi = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} (h_2 h_3 A_1) + \frac{\partial}{\partial u_2} (h_3 h_1 A_2) + \frac{\partial}{\partial u_3} (h_1 h_2 A_3) \right] \\ \nabla \times \mathbf{A} &= \text{curl } \mathbf{A} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 e_1 & h_2 e_2 & h_3 e_3 \\ \frac{\partial}{\partial u_1} & \frac{\partial}{\partial u_2} & \frac{\partial}{\partial u_3} \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix}\end{aligned}\quad (7)$$

In rectangular coordinates (e_1, e_2, e_3) are simply $(\hat{i}, \hat{j}, \hat{k})$ and (u_1, u_2, u_3) are simply (x, y, z) .

Tensors

We are most familiar with a three dimensional space because we can visualise it—and because, as all indications suggest, we live in it. But, as explained earlier, just because we cannot visualise N dimensions it does not mean so many dimensions do not exist: nothing physically or mathematically restricts space to three dimensions; we can freely go all the way up to N dimensions and the laws of physics will still remain valid as governed by mathematics⁶. That is on the one hand; on the other, no law of physics should depend on the frame within which it is described. Using tensors helps us overcome both these problems at once⁷.

We represent N dimensional systems as (x^1, x^2, \dots, x^N) where $1, 2, \dots, N$ are **subscripts** and not indices/exponents/powers. In three dimensions, $N = 3$, we only have (x^1, x^2, x^3) which could stand for the three dimensions of any of our usual Cartesian, spherical or cylindrical coordinate systems. We could then write $\mathbf{A} = A_1 x^1 + A_2 x^2 + A_3 x^3$ or, better still, the N -dimensional form $\mathbf{A} = A_1 x^1 + A_2 x^2 + \dots + A_N x^N$ as either $\sum_i A_i x^i$ or even more compactly, using the summation convention⁸, as $A_i x^i$.

Covariant and contravariant tensors

Consider the system (x^1, x^2, \dots, x^N) in which some quantities A^1, A^2, \dots, A^N are well-defined. Consider another system $(x'^1, x'^2, \dots, x'^N)$ in which some other quantities A'^1, A'^2, \dots, A'^N are well-defined. Suppose these sets of quantities are related as

$$A'^p = \frac{\partial x'^p}{\partial x^q} A^q$$

we call them contravariant tensors of the first order or first rank⁹ or simply **contravariant vectors**¹⁰.

We can extend this to N^2 quantities A^{mn} defined in the same (x^1, x^2, \dots, x^N) system related to another N^2 quantities A'^{rs} defined in the same $(x'^1, x'^2, \dots, x'^N)$ system as above as follows:

$$A'^{rs} = \frac{\partial x'^r}{\partial x^m} \frac{\partial x'^s}{\partial x^n} A^{mn} \quad (8)$$

⁶ As opposed to the laws of physics adhering to our common sense or imagination.

⁷ Although we never really faced a problem since we have long been describing N dimensional systems mathematically, the use of tensors generalises our work in an efficient and compact manner.

⁸ The summation convention tells us that repeating indices (also called 'dummy indices') must be summed over from 1 to N while non-repeating indices hold the same index value. Try expanding $c_i a_j \alpha^i$ for $N = 5$.

⁹ 'Rank' or 'order' refers to the number of indices a symbol has.

¹⁰ A vector is a tensor of rank one.

This is a **contravariant tensor** of rank two.

We can similarly define covariant vectors or go straight to **covariant tensors** as follows:

$$A'_{rs} = \frac{\partial x^m}{\partial x'^r} \frac{\partial x^n}{\partial x'^s} A_{mn} \quad (9)$$

A tensor need not be purely either covariant or contravariant; a tensor with both **upper and lower indices** may exist:

$$A'^r_s = \frac{\partial x'^r}{\partial x^m} \frac{\partial x^n}{\partial x'^s} A^m_n \quad (10)$$

We call these **mixed tensors**. A famous mixed tensor is the Kronecker delta. Defined as

$$\delta^j_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

the Kronecker delta is a second order mixed tensor.

10 things to keep in mind about tensors¹¹

Treat this as a quick overview of tensors, or a handy guide to refer back to when you need to, at least until you are fully familiar with tensorial mathematics.

1. *Four vector transformation.* $V^\alpha \rightarrow V'^\alpha = \Lambda^\alpha_\beta V^\beta$ or $V_\alpha \rightarrow V'_\alpha = \Lambda^\beta_\alpha V_\beta$
2. *Lorentz transformation.* For $\partial x^\alpha / \partial x'^\beta = \Lambda^\alpha_\beta$ we have $\Lambda^\alpha_\beta = \eta_{\alpha\sigma} \eta^{\beta\rho} \Lambda^\sigma_\rho$
3. *The Kronecker delta function.* $\Lambda^\alpha_\beta \Lambda^\beta_\rho = \delta^\alpha_\rho = 1$ if $\alpha = \rho$, else $= 0$.
4. *Scalar product invariance.* $A'_\alpha B'^\alpha = \Lambda^\alpha_\beta \Lambda^\rho_\alpha A_\beta B^\rho = \delta^\beta_\rho A_\beta B^\rho = A_\beta B^\beta$
5. *Raising and lowering indices.* $A_\alpha = \eta_{\alpha\beta} A^\beta$ and $A^\alpha = \eta^{\alpha\beta} A_\beta$
6. *Linearly combinable.* $A^\alpha_\beta = m B^\alpha_\beta + n C^\alpha_\beta$
7. *Direct product.* $A^{\alpha\rho}_\beta = B^\alpha_\beta C^\rho$
8. *Contraction.* $A^{\alpha\gamma}_\beta = A^{\alpha\gamma\delta}_\beta$ if we set upper index $\delta =$ lower index β .
9. *Differentiation.* $A^{\alpha\beta}_\gamma = \partial_\gamma x^\alpha A^{\alpha\beta}$ where the notation $\partial_\gamma x^\alpha := \partial / \partial x^\gamma$.
10. *Levi-Civita tensor.* $\epsilon^{\alpha\beta\gamma\delta}$ defined to be $+1$ for even permutations¹² of $(\alpha, \beta, \gamma, \delta)$, -1 for odd permutations, and zero in all other cases.

Generally, it is also worth remembering, we say two tensors are equal if they have the same covariant rank, the same contravariant rank and equal corresponding components.

Also, a tensor $A^{\alpha\beta}$ has a symmetric counterpart in $A^{\beta\alpha}$ if the two are equal, i.e. $A^{\alpha\beta} = A^{\beta\alpha}$. On the same lines it has an antisymmetric component if the two are related as $A^{\alpha\beta} = -A^{\beta\alpha}$. This holds for covariant tensors too.

¹¹ To show that a relationship yields a tensor consider the primed version of that relationship and convert it into the unprimed form: you must end up with precisely the same equation but with no prime and with different indices obeying the Lorentz transformation described in 2 above. As an exercise use this technique to prove that the left-hand sides of points 6, 7, 8 and 9 listed on the right are in fact tensors.

¹² An even permutation means an even number of swaps. For example $(\gamma, \delta, \alpha, \beta)$ is an even permutation of $(\alpha, \beta, \gamma, \delta)$ since it has two swaps, $\alpha \leftrightarrow \gamma$ and $\beta \leftrightarrow \delta$. Likewise an odd permutation has an odd number of swaps, e.g. with one swap $\beta \leftrightarrow \gamma$ you can get $(\alpha, \gamma, \beta, \delta)$.

Determining if a quantity is a tensor—the quotient law

Before we head to the next major discussion in this topic it is worth taking a look at the quotient law, a handy method we can use to determine if a quantity that merely ‘looks’ like a tensor is, in fact, a tensor.

First, not all Q^α_β are tensors. We can find out if a quantity is really a tensor by using the inner product operation. More specifically, *if the inner product between a questionable quantity X and a known tensor is a tensor then X is also a tensor.*

Consider the following example. Say we have $X := X^\lambda_{\mu\nu}$ of whose tensorial nature we are unaware. Let us compute its inner product with a known tensor, say, A^μ and see if we end up with a tensor:

$$A^\mu P^\lambda_{\mu\nu} = \frac{\partial x_\mu}{\partial x'_\gamma} \frac{\partial x_\lambda}{\partial x'_\alpha} \frac{\partial x'_\gamma}{\partial x_\mu} \frac{\partial x'_\alpha}{\partial x_\nu} A'^\gamma P'^\alpha_{\gamma\beta}$$

And there we have it: $P'^\alpha_{\gamma\beta}$ is a third order mixed tensor.

Christoffel symbols

All of our discussions so far have led us to the doorstep of an important topic that will be central to our discussions on General Relativity. First, keep in mind that Christoffel symbols¹³ are not tensors. Second, they come in two types and their core purpose is to describe parallel transport on a metric¹⁴.

The **Christoffel symbol of the first kind** is defined as

$$\Gamma_{\mu\nu,\sigma} := [\mu\nu,\sigma] = \frac{1}{2} \left(\frac{\partial g_{\mu\sigma}}{\partial x_\nu} + \frac{\partial g_{\nu\sigma}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\sigma} \right) \quad (11)$$

and the **Christoffel symbol of the second kind** is

$$\Gamma^\sigma_{\mu\nu} := \{\mu\nu,\sigma\} = \frac{1}{2} g^{\sigma\lambda} \left(\frac{\partial g_{\mu\lambda}}{\partial x_\nu} + \frac{\partial g_{\nu\lambda}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\lambda} \right) \quad (12)$$

We quickly notice a couple of relations here. For starters we have $\Gamma_{\mu\nu,\sigma} = \Gamma_{\nu\mu,\sigma}$ or, in the simpler notation given above,

$$[\mu\nu,\sigma] = [\nu\mu,\sigma]$$

Similarly, $\Gamma^\sigma_{\mu\nu} = \Gamma^\sigma_{\nu\mu}$ or, equivalently,

$$\{\mu\nu,\sigma\} = \{\nu\mu,\sigma\}$$

meaning Christoffel symbols are symmetric with respect to μ and ν .

Further, there are fairly straightforward relations to go from eq. (12) to (11) and vice versa. These are given by

$$\{\mu\nu,\sigma\} = g^{\sigma\lambda} [\mu\nu,\lambda] \quad \text{and} \quad [\mu\nu,\sigma] = g_{\sigma\lambda} \{\mu\nu,\lambda\}$$

Try to prove these results¹⁵.

To end with we shall take note of one final result: The following specific sum of two Christoffel symbols of the first kind

$$[\mu\nu,\sigma] + [\sigma\nu,\mu] = \frac{\partial g_{\mu\nu}}{\partial x_\sigma}$$

¹³ Sometimes also Christoffel’s symbols.

¹⁴ Do not worry about these ideas just yet: they will be clarified once you discuss GR at some length.

¹⁵ Simply consider eq. (11) for starters, replace σ by λ and then multiply by $g^{\sigma\lambda}$ on both sides. The reverse is left to you to figure out along similar lines.

Problems

To round off our discussion on tensors let us look at a couple of problems that should put our abstract ideas of the metric g and the Christoffel symbols Γ on a firmer footing¹⁶.

Problem 1 (To see how a metric describes surfaces). *Find the metric tensor that describes the surface of a sphere.*

Solution. The surface we are looking at (that of a sphere) is intrinsically two-dimensional¹⁷. Let its radius be some r . We know then that its surface is given by the general equation

$$ds^2 = g_{ij}e^i e^j$$

in two dimensions. And we also know that in our case $e^i = r d\theta$ and $e^j = r \sin \theta d\phi$ (we would have had another $e^k = dr$ had we been in three dimensions). On substituting,

$$ds^2 = r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

where $g_{11} = r^2$, $g_{22} = r^2 \sin^2 \theta$, the coefficients of the coordinates, and the rest are zeroes:

$$g = \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} = \begin{vmatrix} r^2 & 0 \\ 0 & r^2 \sin^2 \theta \end{vmatrix} = r^4 \sin^2 \theta$$

We now proceed to find $g^{\alpha\beta}$, the metric tensor components, using the above result. Specifically, note that this is given by the equation

$$g^{\alpha\beta} = \frac{\text{cofactor}(g_{\alpha\beta})}{g}$$

so that e.g. $e^{11} = r^2 \sin^2 \theta / r^4 \sin^2 \theta = r^{-2}$ and so on. All four terms turn out as follows:

$$g^{11} = r^{-2} \quad , \quad g^{22} = r^{-2} \sin^{-2} \theta \quad \text{and} \quad g^{12} = 0 = g^{21}$$

□

Problem 2 (Finding Christoffel symbols). *Find the Christoffel symbols corresponding to the equation for ds^2 in the previous problem.*

Solution. From the previous problem we know the values of $g_{\alpha\beta}$ so we can proceed without systematically arriving at those again. We begin with eq. (11) setting $\mu = 2$ and $\nu = 2$ and $\sigma = 1$ to get¹⁸

$$\begin{aligned} [22, 1] &= \frac{1}{2} \left(\frac{\partial g_{21}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1} \right) \\ &= \frac{1}{2} \left(0 + 0 - \frac{\partial}{\partial \theta} (r^2 \sin^2 \theta) \right) \\ &= -a^2 \sin \theta \cos \theta \end{aligned}$$

Finding the rest—namely $[12, 2]$, $\{22, 1\}$ and $\{12, 2\}$ —is left to you as an exercise¹⁹. When finding the Christoffel symbols of the second kind be sure to substitute the contravariant and covariant forms of g appropriately.

□

¹⁶ The two problems solved here are in two dimensions. Review them and then try solving them in three dimensions.

¹⁷ Intrinsic dimensions are what one observes a surface to be when one is on that surface, e.g. an ant stuck to the surface of a sphere thinks the surface is two dimensional; to see the third dimension as we do the ant would have to leave the surface of the sphere—this is called the extrinsic dimension of the sphere and it is in this sense that the surface becomes three-dimensional.

¹⁸ Remember that $g_{\alpha\beta} \neq g^{\alpha\beta}$ so substitute accordingly, carefully. The covariant form is obtained without the need for the determinant g .

¹⁹ All others will turn out to be zeroes.