3.4 Lectures 4 - 5: Tensors and the effects of gravity

3.4.1 Tensors: A brief introduction

In this section we will explain the concept and use of tensors in as simple terms as possible. The background assumed here is classical mechanics according to Newton, quantum mechanics at the first year master level and special relativity. The application of tensors in general relativity, or any other context, will then follow as another example of the cases discussed below.

What is a tensor? Consider the following steps
1. Specify a set of symmetry transformations (i.e., a group G)
e.g. rotations in 3d (G = SO(3), or SU(2) in QM), or SO(1,3) for Lorentz rotations in Minkowski space,
2. Select transformation prototypes: objects with one index ("vectors")
e.g., 3d vectors $\mathbf{r}^T = (x, y, z) = x^i$, complex 2-component spinors χ_a in QM, 4-vectors V^α in Mink space; we have then: $\mathbf{r}' = R\mathbf{r}$, $\chi' = g\chi$ and $V' = \Lambda V$.
3. Generalise the prototypes to any tensor: objects with more than one index
e.g., the (symmetric) stress tensor $T_{\alpha\beta}$, the (antisymmetric) EM field strength $F_{\alpha\beta}$.

Definition: A tensor is a multi-indexed object like $T_{ijk...}$ or $T_{\alpha....\beta}{}^{\gamma...\delta}$ which transforms under its respective group of transformations G as follows: Each index rotates in the same way as the corresponding prototype does.
Ex 1: Since $x'^i = R^i{}_j x^j$ we have $x'_i = R_i{}^j x_j$ and thus also $T_{ijk} \to T'_{ijk} = R_i{}^l R_j{}^m R_k{}^n T_{lmn}$. Note that since these indices are raised and lowered by δ^{ij} and δ_{ij} their position, upper or lower, does not matter.
Ex 2: Since $V'^\alpha = \Lambda^\alpha{}_\beta V^\beta$ we have also $V'_\alpha = \Lambda_\alpha{}^\beta V_\beta$ where $\Lambda_\alpha{}^\beta := \eta_{\alpha\gamma}\Lambda^\gamma{}_\delta \eta^{\delta\beta}$, and hence $T'_{\alpha\beta}{}^\gamma = \Lambda_\alpha{}^\delta \Lambda_\beta{}^\epsilon \Lambda^\gamma{}_\zeta T_{\delta\epsilon}{}^\zeta$. Note that since $V^\alpha V_\alpha$ is invariant it follows that $\Lambda^\alpha{}_\beta \Lambda_\alpha{}^\gamma = \delta^\gamma_\beta$ or $\Lambda^\alpha{}_\beta \eta_{\alpha\delta} \Lambda^\delta{}_\epsilon \eta^{\epsilon\gamma} = \delta^\gamma_\beta$ which is the standard relation for a Lorentz transformation $\eta_{\alpha\delta}\Lambda^\alpha{}_\beta \Lambda^\delta{}_\epsilon = \eta^{\beta\epsilon}$.

Note: There are special tensors which stay numerically the same under a transformation, called *invariant tensors*. Examples are δ_{ij} in 3d, ϵ_{ab} and Pauli matrices $(\sigma^i)_a{}^b$ in SU(2), $\eta_{\alpha\beta}$ and $\epsilon^{\alpha\beta\gamma\delta}$ in Mink space. This can never happen for vectors. If it does then the symmetry is broken down to the subset of symmetries that leave the vector intact (e.g., the vector (1,0,0) in 3d breaks SO(3) to SO(2), the rotations around the vector (1,0,0)).

Ex 3: We have been careful in the examples above to use different types of indices for different symmetry transformations and for different kinds of representations. An example of the latter from QM is the rotation properties possessed by a state of two electrons $|\chi_a \chi_b>$. The rules we learn in undergraduate QM is that the answer is one spin 1 state and one spin 0 state. Mathematically this is written as

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1, \tag{3.135}$$

where $\otimes$ denotes a tensor product and $\oplus$ a sum of tensors. This tensor product rule is proven easily by noting that any 2 by 2 matrix, which $|\chi_a\chi_b>$ is an example of, can be expanded in terms of the two kinds of SU(2) invariant tensors, the antisymmetric ϵ_{ab} and the three symmetric Pauli matrices defined by $\sigma^i_{ab} := \sigma^{i\ c}_{a}\epsilon_{cb}$. Thus

$$|\chi_a\chi_b>= -\frac{1}{2}\epsilon_{ab}(\epsilon^{cd}|\chi_c\chi_d>) + \frac{1}{2}\sigma^i_{ab}(\sigma^{cd}_i|\chi_c\chi_d>). \tag{3.136}$$

Here ϵ^{ab} is the inverse of ϵ_{ab}, i.e., $\epsilon_{ab}\epsilon^{bc} = \delta^c_a$ and thus $\epsilon_{ab}\epsilon^{ab} = -2$. All spinor indices are raised and lowered by acting with ϵ from the left: $\chi^a := \epsilon^{ab}\chi_b$ and $\chi_a := \epsilon_{ab}\chi^b$. The expansion coefficients given here in terms of the invariant matrices ϵ and σ^i are called Clebsch-Gordan coefficients: The first term is the spin 0 term (or scalar) and the second is the spin 1 term (or vector) as seen by the index structure of the expressions in the brackets containing $|\chi_c\chi_d>$. In general, for SU(2), which is the relevant transformation group here, the whole representation theory can be expressed in terms of multi-indexed objects with only symmetric indices $|\chi_{(a_1}\chi_{a_2}...\chi_{a_n)}>$ with spin $s = \frac{n}{2}$.

Exercise: Verify that the number of degrees of freedom (=number of components) of $|\chi_{(a_1}\chi_{a_2}...\chi_{a_n)}>$ is the same as for spin $s = \frac{n}{2}$.

The three points listed above defines what a tensor is in general. The necessity to introduce tensors in physics stems from the fact that the physical outcome of any experiment cannot depend on the orientation in space or spacetime, neither on the location in space or spacetime as defined in any given coordinate system. Furthermore, different experimentalists must be able to compare their results for identical experiments just differing by which coordinate system they have chosen to describe the results in. Thus the equations used, like e.g. Newton's equations, must have the same form in all coordinate systems and have well-defined transformation properties when comparing them in different coordinate systems.

Consider Newton's 2nd law: $\mathbf{F} = m\mathbf{a}$ or $F^i = ma^i$. As written, it does not refer to any particular coordinate system. In fact, it has this form in any Galilean inertial frame (all frames related by space rotations and translations in space and time). However, if comparing the application of this equation in two different inertial frames related by a space rotation, defined by $x'^i = R^i{}_jx^j$ for some 3 by 3 rotation matrix R, relative each other we must have that in the primed system $F'^i = ma'^i$ where $F'^i = R^i{}_jF^j$ and $a'^i = R^i{}_ja^j$ so that the rotation matrix can be dropped on both sides and we thus find the unprimed equation $F^i = ma^i$. We then say that Newton's equation transforms as a tensor equation, the tensors here being just vectors which are defined to behave under rotations in exactly the same way as the prototype vector x^i.

The next insight comes from scalar products which are scalars (e.g., objects invariant under the transformations in question) constructed by taking (scalar) products of vectors. Examples are $r^2 = x^ix^i = x^ix^j\delta_{ij} = x^ix_i$, $VW = V^\alpha W^\beta\eta_{\alpha\beta} = V^\alpha W_\alpha$ or for spinors in

QM $< \chi'|\chi >:= \bar{\chi}'^{a}\chi_a := \epsilon^{ab}\bar{\chi}'_b\chi_a$. These scalars are invariant under, respectively, SO(3), SO(1,3) and SU(2). This invariance can be shown as follows.

Consider first $r^2 = x^i x^i = x^i x^j \delta_{ij} = x^i x_i$. Rotating x^i by $x'^i = R^i{}_j x^j$ we get that $x^i x^j \delta_{ij}$ becomes $x'^i x'^j \delta_{ij} = R^i{}_k x^k R^j{}_l x^l \delta_{ij} = \mathbf{r}^T R^T \mathbf{1} R \mathbf{r}$ which must equal $\mathbf{r}^T \mathbf{r}$ which means that

$$R^T R = \mathbf{1}, \text{ or } \delta_{ij} R^i{}_k R^j{}_l = \delta_{kl}. \tag{3.137}$$

This result says that rotation matrices are *orthogonal* (first version of the condition) or equivalently that rotations leave the tensor δ_{ij} numerically invariant. Note that a general non-invariant two-index tensor transforms under rotations as

$$T_{ij} \to T'_{ij} = T_{kl} R^k{}_i R^l{}_j \text{ or } T' = R^T T R. \tag{3.138}$$

Note that for tensors with more than two indices the second way of writing it does not exist. This is why in physics we tend to always use index notation for tensors.

In the SU(2) case it is a bit more intricate. The scalar product is defined by means of the ϵ tensor. It is invariant under SU(2) by noting its connection to the determinant which is equal to 1 for all $g \in SU(2)$:

$$\epsilon'_{ab} := g_a{}^c g_b{}^d \epsilon_{cd} = (detg)\epsilon_{ab} = \epsilon_{ab}. \tag{3.139}$$

Then the invariance of the SU(2) scalar product works very similar to how it works in special relativity for SO(1,3) with a scalar product defined in terms of $\eta_{\alpha\beta}$.

In physics all equations are tensor equations written in terms of indices which can appear in *only two ways*: either
1) contracted pairwise as for $\beta\gamma$ in $U_\alpha{}^{\beta\gamma} V_{\beta\gamma} = W_\alpha$, or
2) uncontracted (appears once in each term) as for α in $U_\alpha{}^{\beta\gamma} V_{\beta\gamma} = W_\alpha$.
Each contracted index pair behaves as a scalar product and is thus invariant. The equation as a whole is either invariant (has no uncontracted indices) or covariant (has uncontracted indices like in the example used here). A tensor equation cannot have terms with different sets of uncontracted indices!

Finally: The symmetry transformations we discuss here are all implemented by some kind of "rotation" matrices (for SO(3), SU(2), or SO(1,3)) and hence are supposed to belong to well-defined sets, denoted G in general, of matrices. In each case these matrices $g \in G$ have the properties
1. If g_1 and g_2 belong to the set so does $g_3 = g_2 g_1$. This is natural since two rotations performed after each other must equal a single rotation. This is called *closure*, the set is closed under matrix multiplication if all possible rotation matrices is included in the set. Note that in SO(3) we have $R^T R = 1$ which implies $R_3^T R_3 = (R_2 R_1)^T R_2 R_1 = R_1^T R_2^T R_2 R_1 = 1$ as claimed. The "S" in SO(3) means that $detR = 1$ which is also preserved under matrix multiplication.
2. The set of matrices G obviously contains the unit matrix.
3. There is a unique inverse to each matrix in the set. This is also quite obvious here.
4. The multiplication is associative. This is again clear since associativity is satisfied by all matrices.

The axioms listed in 1) to 4) defines a "group" and as we have seen they are trivially satisfied for matrix groups.

Note: It is often very convenient to define a matrix group by which tensors it leaves invariant. This is of course equivalent to specifying the scalar product under consideration. There is therefore a natural hierarchy of groups. Consider all matrices g that can act on an N-dimensional real vector space. The only requirement for these coming from the four axioms is that $detg \neq 0$ so that its inverse can be defined. This group is called the *general linear group* denoted $GL(N)$. If it leaves invariant the ϵ tensor it must have unit determinant and then it is called *special linear group* denoted $SL(N)$. In these cases, if a vector v^i transforms as $v'^i = g^i{}_j v^j$ then in order to define a scalar (which is not a scalar product since it requires a metric) one must introduce also a covector v_i transforming as $v'_i = v_j (g^{-1})^j{}_i$. The scalar product $v_i v^i$ is then trivially invariant. This is exactly what happens in general relativity in the context of general coordinate transformations.

General coordinate transformations (or diffeomorphisms)
The transformations and prototypes dx^μ and ∂_μ are given by the chain rule

$$x^\mu \to x'^\mu = x'^\mu(x) \Rightarrow dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu, \ \partial'_\mu = \frac{\partial x^\mu}{\partial x'^\nu} \partial x^\nu \Rightarrow d' := dx'^\mu \partial'_\mu = dx^\mu \partial_\mu :_d, \quad (3.140)$$

i.e., the exterior derivative $d = dx^\mu \partial_\mu$ is the basic invariant quantity in this case. This object is of central importance in "differential geometry", a subject we will not discuss further in this course.

There are only two invariant objects in GR, i.e., under general coordinate transformations. Note that the metric itself, contrary to in special relativity, is not an invariant. However, if we consider integrals we know

$$d^4 x' = |\frac{\partial x'^\mu}{\partial x^\nu}| d^4 x, \quad (3.141)$$

where $|\frac{\partial x'^\mu}{\partial x^\nu}|$ is the determinant of $\frac{\partial x'^\mu}{\partial x^\nu}$ called the Jacobian J, i.e., $J = \det(\frac{\partial x'^\mu}{\partial x^\nu})$. Although d^4x has no uncontracted indices it does not behave entirely as a scalar due to the Jacobian. To keep track of such *extra* factors of the Jacobian we use the name **tensor density** of weight $w \in \mathbf{Z}$ for any tensor that transforms according to its indices but with additional w Jacobian factors. Another example is the determinant of the metric:

$$g'_{\mu\nu} = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial x^\mu}{\partial x'^\nu} g_{\rho\sigma} \Rightarrow \det g' = (\det(\frac{\partial x^\mu}{\partial x'^\nu}))^2 \det g, \quad (3.142)$$

which means that $\det g$ has weight $w = -2$ since $\det(\frac{\partial x^\mu}{\partial x'^\nu}) = (\det(\frac{\partial x'^\mu}{\partial x^\nu}))^{-1}$. The importance of these facts are clear when considering integrations since we see that $d^4x\sqrt{-g}$ is an invariant integration measure or volume element dV, i.e., it is a scalar with weight zero. Here $g = \det g_{\mu\nu}$. Note that Weinberg defines $g = -\det g_{\mu\nu}$ which is not so common in the literature! Also quite common is to use an absolute value sign $\sqrt{|g|}$.

Another important object is the totally antisymmetric **Levi-Civita symbol** $\varepsilon^{\mu\nu\rho\sigma}$ which is a tensor with non-zero weight, that is, a tensor density. It is numerically invariant under coordinate transformations and take the following values in all coordinates:

$$\varepsilon^{\mu\nu\rho\sigma}: \ \varepsilon^{0123} = +1. \tag{3.143}$$

To define this object we first consider the related $w = 0$ tensor $\epsilon^{\mu\nu\rho\sigma}$ (note the change in notation from ε to ϵ). It can be obtained from the Minkowski equivalent $\epsilon^{\alpha\beta\gamma\delta}$ as follows:

$$\epsilon^{\mu\nu\rho\sigma} := \frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial x^\nu}{\partial \xi^\beta}\frac{\partial x^\rho}{\partial \xi^\gamma}\frac{\partial x^\sigma}{\partial \xi^\sigma}\epsilon^{\alpha\beta\gamma\delta}, \tag{3.144}$$

and it clearly transforms as (where we distinguish indices by the bar notation)

$$\epsilon'^{\mu\nu\rho\sigma} := \frac{\partial x'^\mu}{\partial x^{\bar{\mu}}}\frac{\partial x'^\nu}{\partial x^{\bar{\nu}}}\frac{\partial x'^\rho}{\partial x^{\bar{\rho}}}\frac{\partial x'^\sigma}{\partial x^{\bar{\sigma}}}\epsilon^{\bar{\mu}\bar{\nu}\bar{\rho}\bar{\sigma}} = \det(\frac{\partial x'^\mu}{\partial x^{\bar{\mu}}})\epsilon^{\mu\nu\rho\sigma}, \tag{3.145}$$

where we have used the definition of a determinant in the last step. Although there is a factor of the Jacobian $J = \det(\frac{\partial x'^\mu}{\partial x^{\bar{\mu}}})$ in this formula it is tied to the index structure so is not mean that this object is a density.

However, what the presence of the Jacobian above does mean is that by multiplying the tensor $\epsilon^{\alpha\beta\gamma\delta}$ by $\sqrt{-g}$ we have obtained a different object, a weight $w = -1$ tensor density denoted $\varepsilon^{\mu\nu\rho\sigma}$ and called the **Levi-Civita symbol**, which is a numerical invariant! That is[9]

$$\varepsilon^{\mu\nu\rho\sigma} = \sqrt{-g}\epsilon^{\mu\nu\rho\sigma}. \tag{3.146}$$

Exercise: Is the $\varepsilon_{\mu\nu\rho\sigma}$ also an invariant tensor density? What is its weight?

The affine connection having defined tensors and tensor densities above we will now investigate the affine connection to see how it transforms under general coordinate transformations. Recall the definition obtained from the equation of free particle motion, namely

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0, \ \text{where} \ \Gamma^\mu_{\nu\rho} = \frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial^2 \xi^\alpha}{\partial x^\nu \partial x^\rho}. \tag{3.147}$$

Is $\Gamma^\mu_{\nu\rho}$ a tensor?

To answer this question we perform a diffeo (=general coordinate transformation) $x \to x' = x'(x)$:

$$\Gamma^\mu_{\nu\rho} \to \Gamma'^\mu_{\nu\rho} = \frac{\partial x'^\mu}{\partial \xi^\alpha}\frac{\partial^2 \xi^\alpha}{\partial x'^\nu \partial x'^\rho} = (\frac{\partial x'^\mu}{\partial \xi^\alpha})\frac{\partial}{\partial x'^\nu}(\frac{\partial \xi^\alpha}{\partial x'^\rho}) = (\frac{\partial x'^\mu}{\partial \xi^\alpha})\frac{\partial}{\partial x'^\nu}(\frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial \xi^\alpha}{\partial x^\sigma}). \tag{3.148}$$

[9] Some authors define the invariant object with all indices down instead of up as here. The virtue of our definition is that in differential geometry the integration measure is $dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma = d^4x\varepsilon^{\mu\nu\rho\sigma} = d^4x\sqrt{-g}\epsilon^{\mu\nu\rho\sigma}$.

Here we do the following: when the derivative hits the second factor in the last bracket we rewrite it as $\frac{\partial}{\partial x'^\nu} = \frac{\partial x^\eta}{\partial x'^\nu}\frac{\partial}{\partial x^\eta}$ and the first bracket we write as $\frac{\partial x'^\mu}{\partial \xi^\alpha} = \frac{\partial x'^\mu}{\partial x^\tau}\frac{\partial x^\tau}{\partial \xi^\alpha}$. When the derivative hits the first factor in the last bracket we keep everything as it is. This gives

$$\Gamma'^\mu_{\nu\rho} = \frac{\partial x'^\mu}{\partial x^\tau}\frac{\partial x^\eta}{\partial x'^\nu}\frac{\partial x^\sigma}{\partial x'^\rho}\Gamma^\tau_{\eta\sigma} + \frac{\partial x'^\mu}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\rho}. \tag{3.149}$$

This is a fundamental result: the RHS has two terms of which the first one is "good" (homogeneous) showing tensor behaviour while the second one is "bad" (inhomogeneous) destroying the tensor property of $\Gamma^\mu_{\nu\rho}$ which is reflected in the name "connection".

The comparison to EM is clear: also the vector potential is a connection (under gauge transformations) since $A'_\alpha = A_\alpha + \partial_\alpha \Lambda$. In EM we know how to take the next step that is how to turn partial derivatives into $U(1)$ covariant ones using the vector potential:

$$D_\alpha \psi := (\partial_\alpha - ieA_\alpha)\psi \Rightarrow \text{if } \psi'(x) = e^{ie\Lambda(x)}\psi(x) \text{ then also } (D_\alpha\psi(x))' = e^{ie\Lambda(x)}D_\alpha\psi(x). \tag{3.150}$$

That this covariance condition is satisfied is easily checked

$$LHS: \; (D_\alpha\psi(x))' := (\partial_\alpha - ieA'_\alpha)\psi' = e^{ie\Lambda(x)}\partial_\alpha\psi(x) + ie(\partial_\alpha\Lambda)e^{ie\Lambda(x)}\psi(x) - ieA'_\alpha e^{ie\Lambda(x)}\psi(x), \tag{3.151}$$

while the RHS is

$$RHS: \; e^{ie\Lambda(x)}D_\alpha\psi(x) = e^{ie\Lambda(x)}\partial_\alpha\psi(x) - ieA_\alpha e^{ie\Lambda(x)}\psi(x). \tag{3.152}$$

Hence we see that the first terms on both sides cancel and the remaining terms just says that $A'_\alpha = A_\alpha + \partial_\alpha\Lambda$.

In a similar way to EM above we now define a covariant derivative ∇_μ in GR. Acting on a contravariant vector V^μ it is defined by

$$\nabla_\mu V^\nu := \partial_\mu V^\nu + \Gamma^\nu_{\mu\rho}V^\rho. \tag{3.153}$$

The covariance requirement in this case is that $\nabla_\mu V^\nu$ should transform as the index structure suggests, i.e., as

$$(\nabla_\mu V^\nu)' = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\sigma}(\nabla_\rho V^\sigma), \tag{3.154}$$

which can verified explicitly once the non-tensor behaviour of the affine connection has been written down.

Note that the geodesic equation is really just a covariant derivative acting on the tangent vector of the particle path. If we set $u^\mu = \frac{dx^\mu}{d\tau}$ it reads

$$D_\tau u^\mu := \partial_\tau u^\mu + u^\nu \Gamma^\nu_{\mu\rho}u^\rho = 0. \tag{3.155}$$

If we replace the tangent vector D_τ is acting on by a vector field V^μ defined in all of spacetime then the LHS of the above equation can be rewritten as follows

$$D_\tau V^\mu = \partial_\tau V^\mu + u^\nu\Gamma^\mu_{\nu\rho}V^\rho = u^\nu(\partial_\nu V^\mu + \Gamma^\mu_{\nu\rho}V^\rho) = u^\nu\nabla_\nu V^\mu, \tag{3.156}$$

where we have used $\partial_\tau = u^\nu\partial_\nu$.