

# Physics W89 - Introduction to Mathematical Physics - Summer 2023

## Problem Set - Module 07 - This Module Makes Me Tensor

*Last Update: July 26, 2023*

Student: Yutong Du

Note: In this problem set, I will be using Einstein summation convention, where all indices that are repeated - occurring once as a lower index and once as an upper index - are assumed to be summed over. We will restrict ourselves to *real* vector spaces for all parts of this problem set unless explicitly stated otherwise.

### Problem 7.1 - Introduction to Tensors

*Relevant Videos: An Old Acquaintance - Dual Vectors; Revisiting Old Friends - Vectors and Matrices; Defining Tensors - Maps; Defining Tensors - Components; The Tensor Product; Contraction*

Consider a type- $(p, q)$  tensor  $T$ . We defined this as a *map* or *machine* that eats a set of  $p$  dual vectors and  $q$  vectors to return a scalar in a fully linear fashion. That is, if we replace any of the  $p$  dual vectors or  $q$  vectors with a linear combination, we can “distribute”. For example, a type- $(0, 2)$  tensor  $F$  eats two vectors to return a scalar and satisfies

$$\begin{aligned} F(c_1 \vec{v}_1 + c_2 \vec{v}_2, \vec{w}) &= c_1 F(\vec{v}_1, \vec{w}) + c_2 F(\vec{v}_2, \vec{w}), \\ F(\vec{v}, c_1 \vec{w}_1 + c_2 \vec{w}_2) &= c_1 F(\vec{v}, \vec{w}_1) + c_2 F(\vec{v}, \vec{w}_2). \end{aligned}$$

We usually define particular tensors by how exactly they act on a set of dual vectors and vectors. For calculations involving tensors, it is useful to introduce its *components*, with the components having  $p$  upper indices and  $q$  lower indices  $T_{[e]}^{i_1 \dots i_p}_{j_1 \dots j_q}$ , defined by acting the tensor on basis elements,

$$T_{[e]}^{i_1 \dots i_p}_{j_1 \dots j_q} \equiv T(\hat{e}^{i_1}, \dots, \hat{e}^{i_p}, \hat{e}_{j_1}, \dots, \hat{e}_{j_q}).$$

Consider a tensor  $\delta$  that eats one dual vector and one vector and whose action is to “feed the vector to the dual vector” to return the scalar. That is,

$$\delta(\underline{\alpha}, \vec{v}) \equiv \underline{\alpha}(\vec{v}).$$

(a) What are the type and rank of the tensor  $\delta$ ? What are the components of  $\delta$ ? Determine  $\delta(\underline{\alpha}, \vec{v})$  if  $\underline{\alpha} \doteq (3 \ 5 \ -2)$  and  $\vec{v} \doteq \begin{pmatrix} -7 \\ 4 \\ 6 \end{pmatrix}$ .

[Supplementary Part (Not for Credit): Argue that this definition is indeed linear in both its dual vector slot and its vector slot.]

Answer (highlight to reveal): [You should find that the components form the Kronecker delta symbol  $\delta^i_j$ ]

*Solution:* Since  $\delta$  eats one dual vector and a vector, it's a type  $(1, 1)$  tensor, hence its rank is 2. The components of  $\delta$  are determined by feeding it basis elements:

$$\delta(\hat{e}^i, \hat{e}_j) = \hat{e}^i(\hat{e}_j) = \delta_j^i$$

Given  $\underline{\alpha}$  and  $\vec{v}$ , we have:

$$\delta(\underline{\alpha}, \vec{v}) = \begin{pmatrix} 3 & 5 & -2 \end{pmatrix} \begin{pmatrix} -7 \\ 4 \\ 6 \end{pmatrix} = -21 + 20 - 12 = -13$$

□

Consider a tensor  $\epsilon$  that eats  $n$  vectors (in an  $n$ -dimensional vector space) and returns the matrix determinant of the  $n \times n$  matrix formed by the column vectors,

$$\epsilon(\vec{v}_1, \dots, \vec{v}_n) \equiv \det \left( \begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_n \end{pmatrix}_{[e]} \right) = \begin{vmatrix} (v_1)^1 & \cdots & (v_n)^1 \\ \vdots & \ddots & \vdots \\ (v_1)^n & \cdots & (v_n)^n \end{vmatrix}.$$

(b) What properties of the determinant ensure that this is a tensor? What are the type and rank of  $\epsilon$ ? If we are in 3-dimensional real space (so  $V = \mathbb{R}^3$  and  $n = 3$ ), what are the components of  $\epsilon$ ?

*Solution:* The determinant “eats”  $n$  vectors, and is also fully linear in each input slot (we covered this in an earlier module). These are the only two requirements needed to satisfy the fact that it is a tensor. Since  $\epsilon$  eats  $n$  vectors, it is a type  $(0, n)$  tensor, and hence it is also rank  $n$ . To find the components of  $\epsilon$  in 3D space, we feed it the basis vectors:

$$\epsilon_{ijk} = \epsilon(\hat{e}_i, \hat{e}_j, \hat{e}_k)$$

Firstly, we know that from the properties of the determinant that if there are two identical columns, then the determinant is automatically zero. For us, this means that every index with a repeated component is zero. For those without repeats, we first compute  $\epsilon_{123}$ :

$$\epsilon_{123} = \det(\mathbb{1}) = 1$$

With that, now we look at how many column swaps it's required to get the other indices:

$$\begin{aligned} \epsilon_{132} &= -\epsilon_{123} = -1 \\ \epsilon_{213} &= -\epsilon_{123} = -1 \\ \epsilon_{231} &= \epsilon_{123} = 1 && \text{since two swaps required: } 123 \rightarrow 321 \rightarrow 231 \\ \epsilon_{321} &= -\epsilon_{123} = -1 \\ \epsilon_{312} &= \epsilon_{123} = 1 && \text{since two swaps required: } 123 \rightarrow 321 \rightarrow 312 \end{aligned}$$

□

(c) Determine the type of each of the following tensor combinations/actions and convert to index notation by acting your tensor on basis vectors and dual vectors.<sup>1</sup>

<sup>1</sup>You should also try to guess what the index notation looks like before explicitly finding the answers. As you get more practice you will find it easier and easier to jump between the map and component descriptions!

1. **Not-for-credit example:**  $\underline{\alpha}(\vec{v}) = \alpha_i v^i$ , which is a type-(0,0) object (a scalar).
2.  $\underline{\alpha} \otimes \vec{v}$ , the tensor product of  $\underline{\alpha}$  and  $\vec{v}$ .

*Solution:* This takes a type (1, 0) and type (0, 1) tensor to create a type (1, 1) tensor, whose index representation is  $\alpha_j v^i$ . □

3.  $M(\underline{\alpha}, \vec{v})$ , a linear transformation  $M$  acting on a dual vector  $\underline{\alpha}$  and a vector  $\vec{v}$ .

*Solution:*  $M$  is a type (1, 1) tensor, and since we are acting it exactly on its “diet”, then the result is a scalar, or a type (0, 0) tensor (a scalar). In index notation, it’s written as:

$$M(\underline{\alpha}, \vec{v}) = \alpha_i M_j^i v^j$$

□

4.  $A \otimes B$ , the tensor product of two linear transformations  $A$  and  $B$ .

*Solution:*  $A$  and  $B$  are both linear transformations, hence they are both type (1, 1) tensors. Thus, their tensor product creates a type (2, 2) tensor, which has index representation  $C_{j_1, j_2}^{i_1, i_2}$ . □

5. **Not-for-credit example:** Feeding a single vector  $\vec{v}$  to a linear transformation,  $M\vec{v} \equiv M(\cdot, \vec{v})$ , which is a type-(1,0) object (a vector), with components  $M_j^i v^j$ .
6.  $G$ , a tensor that eats two dual vectors to return a scalar.

*Solution:* Since  $G$  eats two dual vectors it’s a type (2, 0) tensor. As for the index notation, we should have two upper indices:  $G^{ij}$ . □

7.  $G(\underline{\alpha}, \cdot)$ , the tensor from the previous part fed the dual vector  $\underline{\alpha}$  to the first slot.

*Solution:* This tensor now requires only a single dual vector to become a scalar, hence it is a type (1, 0) tensor. In index notation:  $G^j = \alpha_i G^{ij}$ . We need the  $\alpha$  here since the properties of  $G(\underline{\alpha}, \cdot)$  depend on the dual vector  $\underline{\alpha}$  that we feed in. □

8. **Not-for-credit extra part:**  $G(\cdot, \underline{\beta})$ , the tensor from the previous part fed the dual vector  $\underline{\beta}$  to the second slot.

[Note: Detailed solutions for the not-for-credit examples (1) and (5) are given in the problem set supplement.]

In these next two parts we will consider type-(0,2) tensors which are the most common tensorial objects after scalars, vectors, dual vectors, and linear transformations. These tensors eat two vectors to return a scalar. First, let’s restrict ourselves to a 2D vector space  $V$  so that our indices can only take on two values. A type-(0,2) tensor  $F$  on this space will only have  $2^2 = 4$  components  $F_{ij}$ . We can get two different dual vectors out if we feed  $F$  a vector  $\vec{v} \doteq \vec{v}_{[e]} \equiv \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}$  depending on whether we plug  $\vec{v}$  into the first slot or the second of  $F$ ,

$$F(\vec{v}, \cdot) = (F_{ij} v^i) \hat{e}_j \doteq ((F_{11} v^1 + F_{21} v^2) \quad (F_{12} v^1 + F_{22} v^2)), \quad (1)$$

$$F(\cdot, \vec{v}) = (F_{ij} v^j) \hat{e}_i \doteq ((F_{11} v^1 + F_{12} v^2) \quad (F_{21} v^1 + F_{22} v^2)). \quad (2)$$

Our general rule for converting index notation to matrix notation (our grids of numbers) is that upper indices label rows and lower indices label columns. In lecture we discussed one way to interpret an object like the type-(0,2) tensor - letting the two lower indices label two different ‘types’ of column to give us a “row vector of row vectors”,

$$\mathbf{F} \doteq \mathbf{F}_{[e]} \equiv \begin{pmatrix} (F_{11} & F_{21}) & (F_{12} & F_{22}) \end{pmatrix}.$$

While we can build a consistent method of matrix manipulations out of this, it is an uncommon practice. Since type-(0,2) (and type-(2,0)) tensors are common objects, it is worthwhile to review a less natural but more familiar way of dealing with these, which is to allow the first lower index to ‘improperly’ label rows to make a matrix form of a type-(0,2) tensor:

$$\mathbf{F} \doteq F_{[e]}^{(\text{mat})} \equiv \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix}.$$

[Note: Remember that the objects with the subscript ‘[e]’ are the matrix representations/grids of components with respect to a basis.]

(d) Evaluate  $\vec{v}_{[e]}^\top F_{[e]}^{(\text{mat})}$  and  $(F_{[e]}^{(\text{mat})} \vec{v}_{[e]})^\top$ . Show that the results are the row vector from Eqs. 1 and 2, respectively. Finally, show  $\mathbf{F}(\vec{v}, \vec{w}) = \vec{v}_{[e]}^\top F_{[e]}^{(\text{mat})} \vec{w}_{[e]}$ .

*Solution:* Since all objects here are nicely represented as matrices, we can just perform matrix multiplication to show these properties:

$$\vec{v}_{[e]}^\top F_{[e]}^{(\text{mat})} = (v^1 \quad v^2) \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} = (v^1 F_{11} + v^2 F_{21} \quad v^1 F_{12} + v^2 F_{22})$$

which equals equation 1. Now for the other one:

$$(F_{[e]} \vec{v}_{[e]})^\top = \left[ \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \right]^\top = \begin{pmatrix} F_{11}v^1 + F_{12}v^2 \\ F_{21}v^1 + F_{22}v^2 \end{pmatrix}^\top = (F_{11}v^1 + F_{12}v^2 \quad F_{21}v^1 + F_{22}v^2)$$

which is exactly equation 2. Now as to the proof, we first look at  $\mathbf{F}(\vec{v}, \vec{w})$  from a tensor perspective. We know that for a general tensor:

$$T(\underline{\alpha}_1, \dots, \underline{\alpha}_p, \vec{v}_1, \dots, \vec{v}_q) = (\alpha_1)_{i_1} \dots (\alpha_p)_{i_p} T_{j_1, \dots, j_q}^{i_1, \dots, i_p} (v_1)^{j_1} \dots (v_q)^{j_q}$$

applying this property to our tensor, we get that in index notation,  $\mathbf{F}(\vec{v}, \vec{w}) = F_{ij} v^i w^j$ . Now if we look at the other side of this equation from a matrix multiplication point of view, we get:  $\vec{v}^\top F^{\text{mat}} \vec{w} = v^i F_{ij} w^j$ , which is identical to the index equation for  $\mathbf{F}(\vec{v}, \vec{w})$ . Since their index representations are identical, we conclude they must be equal.  $\square$

Consider an **antisymmetric** type-(0,2) tensor  $\mathbf{A}$ . Antisymmetry here means flipping the order that the vectors are fed to  $\mathbf{A}$  negates the result,  $\mathbf{A}(\vec{v}, \vec{w}) = -\mathbf{A}(\vec{w}, \vec{v})$ . Suppose we are working in our 3D vector space, so there are  $3^2 = 9$  total components  $A_{ij}$  of  $\mathbf{A}$  (three choices for  $i$  and three choices for  $j$ ).

(e) What is the result of feeding  $\mathbf{A}$  two copies of the same vector,  $\mathbf{A}(\vec{v}, \vec{v})$ ? Which of the nine components of  $A_{ij}$  must be 0? How many free parameters are needed to fully specify all components of  $\mathbf{A}$ ?

[*Spoilers!* For example, if we knew  $A_{12}$  we would know what  $A_{21}$  would have to be (why?), so one free parameter determines this pair of components.]

*Solution:* If we feed  $\mathbf{A}$  two copies of the same vector, we get the equation:  $\mathbf{A}(\vec{v}, \vec{v}) = -\mathbf{A}(\vec{v}, \vec{v})$  which means that  $\mathbf{A}(\vec{v}, \vec{v}) = 0$ . As for the components, this means that all the components on the diagonal (which are

found by feeding  $A$  two identical basis vectors) will be zero. Written out, this is  $A_{11} = A_{22} = A_{33} = 0$ . Due to the antisymmetry property of  $A$ , we also get  $A_{12} = -A_{21}$ ,  $A_{13} = -A_{31}$  and  $A_{23} = -A_{32}$ , meaning that these values come in pairs. Hence, we require three free parameters (one for each pair) to fully specify the components of  $A$ .  $\square$

---

Consider a type-(1,2) tensor  $A$  with components  $A^i_{jk}$  and a type-(2,0) tensor  $B$  with components  $B^{\ell m}$ . Let  $D = A \otimes B$ , which is a type-(3,2) tensor with components  $A^i_{jk} B^{\ell m}$ .

(f) Determine all possible tensors that can be formed via a single contraction of  $D$ . For each tensor give the type and the components in index notation.

[Note: An example is given in the problem set supplement.]

[Supplementary Part (Not for Credit): Do the same but this time with all possible contractions.]

[Spoilers! There are 6 possible single contractions.]

*Solution:* Written out in index notation, the tensor  $D$  is:

$$D = D^{ilm} \hat{e}^i \otimes \hat{e}^l \otimes \hat{e}^m \otimes \hat{e}_j \otimes \hat{e}_k$$

As for contractions, we pick one of the dual basis vectors and one basis vector to contract, so there are 6 ways we can do this. I'll list them off:

- Contract  $i, j$ :  $\bar{D} = D^{ilm} \hat{e}^l \otimes \hat{e}^m \otimes \hat{e}_k$
- Contract  $i, k$ :  $\bar{D} = D^{ilm} \hat{e}^l \otimes \hat{e}^m \otimes \hat{e}_j$
- Contract  $l, j$ :  $\bar{D} = D^{ilm} \hat{e}^i \otimes \hat{e}^m \otimes \hat{e}_k$
- Contract  $l, k$ :  $\bar{D} = D^{ilm} \hat{e}^i \otimes \hat{e}^m \otimes \hat{e}_j$
- Contract  $m, j$ :  $\bar{D} = D^{ilm} \hat{e}^i \otimes \hat{e}^l \otimes \hat{e}_k$
- Contract  $m, k$ :  $\bar{D} = D^{ilm} \hat{e}^i \otimes \hat{e}^l \otimes \hat{e}_j$

$\square$

---



---

## Problem 7.2 - Tensors and Transformations

*Relevant Videos: Defining Tensors - Transformations; The Tensor Product; Contraction*

Another way of classifying or defining tensors is by how they *transform* under passive or active transformations. If we are changing from an old basis  $\{\hat{e}_i\}$  to a new basis  $\{\hat{f}_i\}$  with the change-of-basis matrix  $C$  defined via  $\hat{f}_j = C^i_j \hat{e}_i$ , then the components of a type- $(p, q)$  tensor transform as

$$T_{[f]}^{i_1 \dots i_p}_{j_1 \dots j_q} = (C^{-1})^{i_1}_{k_1} \dots (C^{-1})^{i_p}_{k_p} T_{[e]}^{k_1 \dots k_p}_{\ell_1 \dots \ell_q} C^{\ell_1}_{j_1} \dots C^{\ell_q}_{j_q}.$$

This way of classifying tensors can be useful for determining what type of tensorial object we are dealing with and is the way tensors are introduced in many texts.

(a) Write down how each of the following tensor components transform. That is, give the components of the tensor with respect to the  $f$ -basis in terms of the components with respect to the  $e$ -basis.

1. **Not-for-credit example:** The components of a dual vector  $\alpha_i$ .
2. The components of a matrix,  $M^i_j$ .

*Solution:* The general definition of the transform is written as:

$$T_{[f]}^{i_1, \dots, i_p}_{j_1, \dots, j_q} = (C^{-1})^{i_1}_{k_1} \dots (C^{-1})^{i_p}_{k_p} T_{[e]}^{k_1, \dots, k_p}_{\ell_1, \dots, \ell_q} C^{\ell_1}_{j_1} \dots C^{\ell_q}_{j_q}$$

I won't be typing this out for every small subpart, but obviously the same applies for all of these. If we just follow this rule for this first one:

$$M^i_{[f]j} = (C^{-1})^i_k M^k_{[e]l} C^l_j$$

□

3. The components of a type-(2,0) tensor,  $K^{ij}$ .

*Solution:* This tensor has two upper indices, so we have:

$$K^{ij}_{[f]} = (C^{-1})^i_l (C^{-1})^j_m K^{lm}_{[e]}$$

□

4. The components of a type-(1,3) tensor, <sup>2</sup>  $B^i_{jkl}$ .

*Solution:* Again, see above:

$$B^i_{[f]jkl} = (C^{-1})^i_m B^m_{nqp} C^n_j C^q_k C^p_l$$

□

[Note: Detailed solutions for the not-for-credit example (1) are given in the problem set supplement.]

<sup>2</sup>One important tensor of this type is the **Riemann tensor** in general relativity, which encodes the curvature of spacetime!

The tensors  $\delta$  and  $\epsilon$  discussed in the previous problems are examples of a special class of tensors called **isotropic**, which means the components are the same after an orthogonal change-of-basis (i.e. a change of basis between two sets of orthonormal bases). Recall that in an orthogonal change-of-basis,  $C^{-1} = C^\top$ .

(b) Show that the components you found for the tensor  $\delta$  are unchanged by an orthogonal change-of-basis (and thus  $\delta$  is an isotropic tensor).

*Solution:* To do this, we note that  $\delta_{[f]j}^i = (C^{-1}\delta_{[e]}C)_j^i$ . Now, we remember that  $\delta$  in the  $e$  basis was  $\delta_j^i$ , or in other words,  $\delta$  was the identity. Therefore:

$$\begin{aligned}\delta_j^i &= (C^{-1}\mathbb{1}C)_j^i \\ &= (C^{-1}C)_j^i \\ &= \mathbb{1} \\ &= \delta_{[e]j}^i\end{aligned}$$

Since  $\delta$  in the  $f$  basis is also the identity, we conclude that  $\delta$  is an isotropic tensor. □

We showed in lecture that the dual basis vectors changed under a change-of-basis following the rule

$$\hat{f}^i = (C^{-1})^i_j \hat{e}^j.$$

(c) **Extra Part** (*Not for Credit*) Try carrying out the proof again on your own! That is, show that this is indeed the new dual basis,  $\hat{f}^i(\hat{f}_j) = \delta_j^i$ . Argue that  $C_j^i = \hat{f}^i(\hat{e}_j)$  and  $(C^{-1})^i_j = \hat{e}^i(\hat{f}_j)$ . Finally, use this to show that the components of a vector transform contravariantly,  $v_{[f]}^i = (C^{-1})^i_j v_{[e]}^j$ .

We claimed that the contraction of a tensor gives another tensor but we haven't shown that the contracted tensor actually *behaves* like a tensor yet. In this part we will see an example of how such a proof would go. Let  $\underline{\alpha}$  be a dual vector and  $K$  be a type-(2,0) tensor. We can take their tensor product to form a type-(2,1) tensor  $\underline{\alpha} \otimes K$ . We can then perform a contraction on the lower index of  $\underline{\alpha}$  with the first upper index of  $K$  to yield a type-(1,0) tensor (a vector).<sup>3</sup> Define  $\vec{v}$  to be the vector resulting contraction, so the components in the "old basis"  $\{\hat{e}_i\}$  are

$$v_{[e]}^j \equiv \alpha_{[e]i} K_{[e]}^{ij}.$$

(d) Show that performing a change of basis on the components of  $\underline{\alpha}$  and  $K$  yields the vector-transformation rule that we would expect for the components of  $\vec{v}$ .

*[Spoilers! Another way of stating this problem: We want to construct  $v_{[f]}^j \equiv \alpha_{[f]i} K_{[f]}^{ij}$  and, by using the known transformation rules for  $\underline{\alpha}$  and  $K$ , show that  $v_{[f]}^j = (C^{-1})^j_k v_{[e]}^k$ , the usual vector component transformation rule. You already found how  $K^{ij}$  transforms in the previous part!]*

*Solution:* We know that this equation for  $v$  is true in any basis, hence  $v_{[f]}^j = \alpha_{[f]i} K_{[f]}^{ij}$ . Now, we expand the

<sup>3</sup>There is another possible contraction with the second upper index of  $K$ , but the proof will look mostly the same in either case

right hand side in terms of how the dual vector  $\alpha$  and  $K$  transform:

$$\begin{aligned}
\tilde{v}_{[f]}^j &= \alpha_{[f]i} K_{[f]}^{ij} = (\alpha_{[e]j} C_i^j) \left( (C^{-1})_m^i (C^{-1})_n^j K_{[e]}^{mn} \right) \\
&= \alpha_{[e]j} \underbrace{C_i^j (C^{-1})_m^i}_{=\mathbb{1}} (C^{-1})_n^j K_{[e]}^{mn} \\
&= \alpha_{[e]j} (C^{-1})_n^j K_{[e]}^{jn} \\
&= (C^{-1})_n^j \underbrace{\alpha_{[e]j} K_{[e]}^{jn}}_{v_{[e]}^n} \\
&= (C^{-1})_n^j v^n
\end{aligned}$$

Combining the first and last lines together, we get:  $\tilde{v}_{[f]}^j = (C^{-1})_n^j v^n$ , which is exactly how a vector transforms, as desired.  $\square$

---

Consider again the type-(0,2) tensor  $F$  on a 2D vector space that we explored in Problem 7.1(d). Recall that  $F_{[e]}^{(\text{mat})}$  is the components of the tensor  $F$  with respect to the old basis arranged in matrix form.

**(e) Extra Part** (*Not for Credit*) Show that under a change-of-basis, the components of  $F$  transform as  $F_{[f]}^{(\text{mat})} = C^\top F_{[e]}^{(\text{mat})} C$ .

*Commentary: If we are transforming between two orthonormal bases then  $C^\top = C^{-1}$  so the transformation law  $\tilde{F}_{[f]} = C^\top F_{[e]}^{(\text{mat})} C$  is the same as the usual matrix transformation law. However, if either of the bases are not orthonormal, then  $C^\top \neq C^{-1}$ . In this case, the matrix form of the transformation of a linear transformation  $M$  is  $M_{[f]} = C^{-1} M_{[e]} C$  but the transformation of the matrix form of a type-(0,2) tensor is still  $F_{[f]}^{(\text{mat})} = C^\top F_{[e]}^{(\text{mat})} C$ .*

---



## Problem 7.3 - Stress<sup>4</sup>

*Relevant Videos: Defining Tensors - Maps; Defining Tensors - Components; Defining Tensors - Transformations; Application - Rotational Dynamics*

In continuum mechanics, **stress** is a physical quantity that expresses the internal forces that neighboring particles of a continuous material exert on each other. Consider a very small cube of jello in static equilibrium. Since our cube is assumed small, any applied force on a face can be assumed to act uniformly over the entire face of the cube.<sup>5</sup> Since our cube is in mechanical equilibrium, any applied force on one face must be balanced by an equal and opposite force on the opposite face of the cube. The stress in this case is a pair of applied forces-per-unit-area on opposite faces of the cube and carries the same units as pressure,  $\text{N/m}^2$ .

A **normal stress** is a force applied normal to a surface, such as pulling the top face up and the bottom face down as shown in the left figure in Fig. 1. A **shear stress** is a force applied parallel to a surface, such as sliding the top face to the right and the bottom face to the left as shown in the right figure in Fig. 1. In all, there are nine possible components of stress corresponding to the three possible directions of force on the three possible pairs of faces.

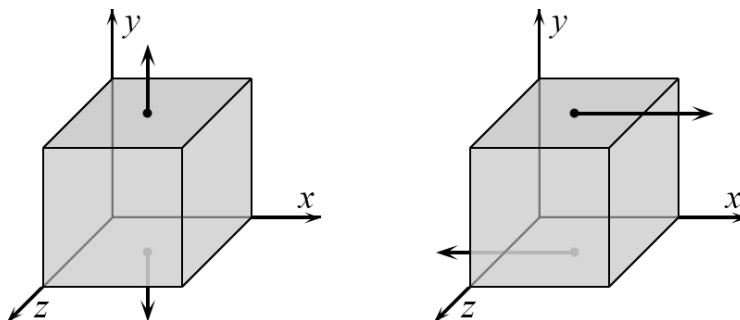


Figure 1: Left: A normal stress on the top/bottom faces. Right: A shear stress on the top/bottom faces.

The **stress tensor**  $\sigma$  is a type-(0,2) tensor whose components are these nine components. That is,  $\sigma(\hat{e}_i, \hat{e}_j)$  gives the  $\hat{e}_j$ -component of the force-per-unit-area acting on the face normal to the direction  $\hat{e}_i$ . For example,  $\sigma(\hat{x}, \hat{z})$  gives the  $z$ -component of the force-per-unit-area acting on the right face (since the normal to the right face is in the  $+\hat{x}$ -direction).

---

**(a) Extra Part (Not for Credit)** Suppose the stress tensor has components  $\sigma_{ij}$ . Express the tensor  $\sigma$  as an expansion in terms of its components and basis vectors/dual vectors/tensor products of basis objects. Let  $\hat{n}$  be a unit vector. Argue why  $\sigma(\hat{n}, \cdot)$  is a dual vector. Express this object in index notation.

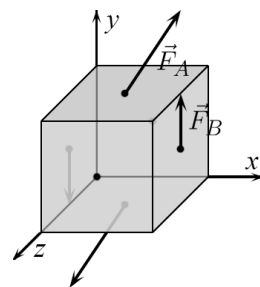
*Commentary: We can interpret  $\sigma(\hat{n}, \cdot)$  as a dual vector  $\underline{F}$  that describes the force-per-unit-area on the cube face for which  $\hat{n}$  is the normal vector. Force is indeed most naturally interpreted as a dual vector since it can be defined via a gradient and gradients are most naturally dual vectors rather than vectors (we will see this when we introduce the Jacobian). Another way of seeing this? The dual vector  $\underline{F}$  eats a displacement vector  $\vec{s}$  to give the work done by the force, a scalar  $W = \underline{F}(\vec{s})$ .*

---

<sup>4</sup>The theory of stresses was studied by Augustin-Louis Cauchy, a French mathematician from the early 1800s. Cauchy introduced the stress tensor that we will study in this problem. Cauchy called these stresses “tensions” and after this, objects which mathematically behaved like these *tens*-ions all got grouped together with one name - tensors!

<sup>5</sup>We don’t “poke the center of the top face” but rather “press down on the entire top face”.

Suppose our cube has uniform side-lengths  $\ell$  and we apply forces as shown to the right to the cube. The forces are  $\underline{F}_A = (2f)\underline{\hat{x}} + (3f)\underline{\hat{y}}$  and  $\underline{F}_B = (2f)\underline{\hat{y}}$ , where  $f$  is some constant with units of force and  $\{\underline{\hat{x}}, \underline{\hat{y}}, \underline{\hat{z}}\}$  are the basis dual vectors.<sup>6</sup>



(b) Find the nine components of the stress tensor,  $\sigma_{ij}$ . Arrange these components in the form of a matrix  $\sigma^{\text{mat}} \equiv \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$  where the first lower index ‘improperly’ labels the row and the second lower index labels the column.

[*Spoilers!* Your answer should be a symmetric matrix.]

*Solution:* To find the components, we feed  $\sigma$  the basis vectors. Just as the problem described,  $\sigma(\underline{\hat{x}}, \underline{\hat{x}})$  refers to the stress placed on the right face acting in the positive  $\underline{\hat{x}}$  direction. Repeating this process for the remaining 8 basis vectors, we get:

$$\sigma = \begin{pmatrix} 0 & 2f/\ell^2 & 0 \\ 2f/\ell^2 & 3f/\ell^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{f}{\ell^2} \begin{pmatrix} 0 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

□

Since we are in static equilibrium, we want the net *torque* on our cube to be zero as well. It turns out that this requirement implies that the stress tensor must be *symmetric*.

(c) **Extra Part** (*Not for Credit*) Show that the net torque for the setup from part (c) is indeed zero (take the pivot point to be the *center* of the cube). Consider a *non*-symmetric stress tensor  $\sigma_{12} = p$  with all other components zero and show that the net torque is non-zero in this case.

We can use the metric to raise the first index of  $\sigma_{ij}$  to create an object  $\tilde{\sigma}$  which is a linear transformation and whose components form a matrix.<sup>7</sup> Since we are dealing with Cartesian coordinates in regular 3D space, it turns out that the matrix representation for  $\tilde{\sigma}$  is the same as the matrix you found part (b). Since we are now dealing with a linear transformation, we can talk about eigenvalues and eigenvectors!

(d) Find the eigenvalues and eigenvectors of  $\tilde{\sigma}$ . The eigenvalues are called the **principal stresses** and the planes normal to the eigenvectors are the **principal planes**.

[*Spoilers!* In part (b) you should have found  $\sigma \doteq \frac{f}{\ell^2} \begin{pmatrix} 0 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ , where the entry in the  $i$ -th row and  $j$ -th column is the component  $\sigma_{ij}$ .]

*Solution:* Building off part (b), we know that the matrix is:

$$\tilde{\sigma} = \frac{f}{\ell^2} \begin{pmatrix} 0 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

<sup>6</sup>We typically work with forces as vectors rather than dual vectors. Really we can consider  $\vec{F}$  as the “raised” version of  $\underline{F}$ , with  $F^i = g^{ij}F_j$ . Since we are typically working with Cartesian coordinates, the components of the metric are the same as the components of the Kronecker delta and thus the vector and dual vector versions of force have the same components.

<sup>7</sup>What this linear transformation does is eat a unit vector and return the *vector version* of the force acting on the plane normal to the unit vector,  $\vec{F} \equiv \tilde{g}(\underline{F})$ .

Solving this matrix for eigenvalues and eigenvectors (I just used Mathematica for this since we've already had our homework on finding eigenvectors and eigenvalues) we get:

$$\lambda_1 = \frac{4f}{\ell^2}, \vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

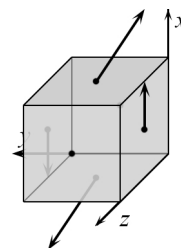
$$\lambda_2 = -\frac{f}{\ell^2}, \vec{v}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda_3 = 0, \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

□

(e) **Extra Part** (*Not for Credit*) Redraw the cube oriented so the six faces are parallel to the principal planes and draw in the forces/stresses on this new picture.<sup>8</sup> Why will analyzing this system with this picture may be more straightforward than analyzing our original system?

(f) Suppose we performed a change-of-basis by rotating our basis vectors 90° counterclockwise about the  $\hat{z}$ -axis, as shown to the right (while keeping the actual stresses unchanged). Based on this new picture, what are the new components of the stress tensor  $\sigma_{[\text{new}]ij}$ ? Show that this is consistent with a passive transformation of your answer from part (b) using the change-of-basis matrix  $C = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .



[*Spoilers!* The result from Problem 7.2(e) will be very useful here!]

*Solution:* Working it out from the diagram using the same logic as we did in part (b), we get that the new stress matrix is:

$$\sigma_{[\text{new}]} = \frac{f}{\ell^2} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We then use the fact that  $\sigma_{[\text{new}]} = C^{-1}\sigma_{[\text{old}]}C$ , and indeed doing this matrix multiplication (done via computer) gives:

$$\sigma_{\text{new}} = \frac{f}{\ell^2} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which does match our desired result.

□

<sup>8</sup>We can redraw the cube because we are really talking about infinitesimal volume elements of our object when we are talking stresses, so we can pick any cube orientation we want.

## 7.4 - Strain and Stiffness

*Relevant Videos: The Tensor Product; Contraction; Application - Rotational Dynamics*

In continuum mechanics, **strain** is the measure of the relative deformation of a small element of material (such as our small cube of jello from the previous problem). For example, if we were considering a 1D spring example of equilibrium length  $\ell$  stretched an additional distance  $\Delta\ell$ , then the strain of the spring would be the ratio  $\Delta\ell/\ell$ .

For our 3D example, let  $x_0$ ,  $y_0$ , and  $z_0$  be the equilibrium lengths of our cube. There are many strains we can describe: We can apply a **normal strain** (a stretch or a compression) by stretching the *right face* to the right by an amount  $\Delta x$ , as shown in the left figure in Fig. 2 (the deformation is in the direction of the normal to the face). The value of this strain is expressed as the dimensionless ratio  $\epsilon_{11} \equiv \Delta x/x_0$ . We can also apply a **shear strain** (a shearing of the cube) by sliding the *top face* of the cube in to the right by an amount  $\Delta x$ , as shown in the right figure in Fig. 2 (the deformation is in a direction perpendicular to the normal to the face). The value of this strain is expressed as the dimensionless ratio  $\epsilon_{12} \equiv \Delta x/y_0$ . There are three possible strains (one normal, two shear) that we can apply to each of the three pairs of faces of the cube, for a total of nine strains. Ooh, this *also* looks like a tensor! The **strain tensor**  $\epsilon$  is a type-(0,2) tensor with components  $\epsilon_{ij}$  defined to eat a vector and return a dual vector.

- The first index again tells you *which face* of the cube you are looking at.
- The second index tells you about the *components of the deformation*.
- Example:  $\epsilon_{23}$  tells you that there is a deformation of the top face (which has normal vector  $\hat{y} = \hat{e}_2$ ) in the direction of  $\hat{z} = \hat{e}_3$ .
- We take the strain tensor to be *symmetric* in its two lower indices,  $\epsilon_{ij} = \epsilon_{ji}$ .

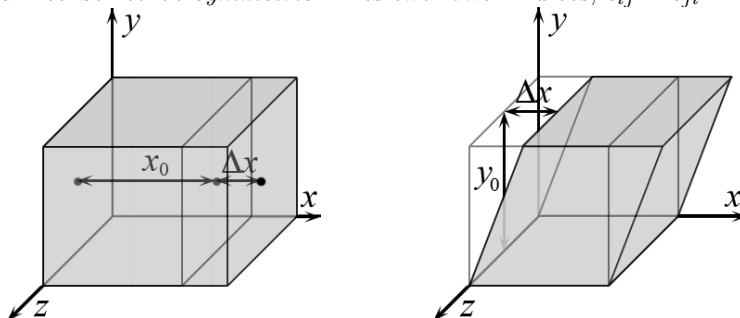


Figure 2: Left: A normal strain of the right/left faces. Right: A shear strain of the top/bottom faces.

---

Consider the strain tensor with components  $\epsilon_{12} = \epsilon_{21} = a$  with all other components zero, where  $a$  is some dimensionless small number.

- (a) **Extra Part** (*Not for Credit*) Describe the shape or sketch a picture of the cube under this strain.
- (b) Write down the components of  $\epsilon$  for the following deformations (you may arrange these components in a matrix,  $\epsilon^{(\text{mat})}$  where component  $\epsilon_{ij}$  is the entry in the  $i$ -th row and  $j$ -th column):
1. **Not-for-credit example:** The left- and right-sides of the cube are compressed inward by a fractional amount  $\delta$ . The rest of the cube remains unchanged.

2. The left- and right-sides of the cube are compressed inward by a fractional amount  $\delta$  and correspondingly the other four faces are bulged out by a fractional amount  $\beta$ .

*Solution:* Just as an example for the logic here, we have  $\epsilon_{11}$  to be the strain in the  $\hat{x}$  direction on the right face, which we have is  $-\delta$ . Doing the same for the other elements, we get:

$$\begin{array}{lll} \epsilon_{11} = -\delta & \epsilon_{12} = 0 & \epsilon_{13} = 0 \\ \epsilon_{21} = 0 & \epsilon_{22} = \beta & \epsilon_{23} = 0 \\ \epsilon_{31} = 0 & \epsilon_{32} = 0 & \epsilon_{33} = \beta \end{array}$$

□

3. The left- and right-sides of the cube are compressed inward by a fractional amount  $\delta$  and sheared by an equal amount upwards/downwards.

*Solution:* Here, we have to account for the fact that the net torque on the whole system is zero, so there's actually another shear on the top and bottom faces to cancel out this torque. Applying the same logic as before now, we get:

$$\begin{array}{lll} \epsilon_{11} = -\delta & \epsilon_{12} = \delta & \epsilon_{13} = 0 \\ \epsilon_{21} = \delta & \epsilon_{22} = 0 & \epsilon_{23} = 0 \\ \epsilon_{31} = 0 & \epsilon_{32} = 0 & \epsilon_{33} = 0 \end{array}$$

□

[Note: Detailed solutions to the not-for-credit example (1) are given in the problem set supplement.]

[Spoilers! Symmetry of the strain tensor means that the described strain in the setup (3) means that what is described isn't the *only* deformation happening to the cube...]

The strain is connected to the stress via a form of **Hooke's law** for elastic materials (in 1D this becomes  $F = -k\Delta x$ ). To lowest order, this will be a linear relationship, so we have

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\epsilon},$$

where  $\mathbf{C}$  is the **stiffness tensor** describing the various elastic properties of the material. The combination  $\mathbf{C}\boldsymbol{\epsilon}$  indicates some form of tensor-product-and-contractions between the stiffness and strain tensors.

(c) Given that  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\sigma}$  are both type-(0,2) tensors, determine the possibilities for the tensor types of  $\mathbf{C}$ . Write out what Hooke's law would look like in *component form* for each of these cases. What units must the components of stiffness possess?

*Solution:* Firstly, we note that any contraction within  $C$  effectively means nothing, so this means that any contraction that occurs must be between a basis vector of  $C$  and one of  $\epsilon$ . This means that for every basis vector we contract, we have to add another so that the balance of a (0,2) tensor on both sides is restored. Ultimately, this means that there are three possibilities for  $C$ .  $C$  can either be a type (0, 0) (a scalar), a type (1, 1), or a type(2, 2). Writing these out:

- Type (0, 0):  $\sigma_{ij} = C\epsilon_{ij}$ .
- Type(1, 1):  $\sigma_{ij} = C_i^l \epsilon_{lj}$

- Type(2, 2):  $\sigma_{ij} = C_{ij}^{lm} \epsilon_{lm}$

In terms of the units, we know that  $\sigma$  is dimensionless and  $\epsilon$  is also dimensionless, so this means that  $C$  must carry all the dimension, hence  $C$  must have units of force/area.  $\square$

---

The inverse of Hooke's law may be written  $\epsilon = \mathbf{S}\sigma$ , where  $\mathbf{S}$  is known as the *compliance tensor* and is of the same type as  $\mathbf{C}$ . Let's design a physical experiment in the lab to narrow down the choices of tensor type we found in the previous part: we will put our cube in a vice and apply an inward stress/pressure on the left- and right-faces. That is,  $\sigma_{11} = -P$  with all other components zero.

(d) For each of the three Hooke's laws from the previous part, determine which components of strain can be non-zero when the described stress is applied. Intuitively, how do you think an elastic object would respond to this strain? Use this to determine which of the three laws is most appropriate and why.

*Solution:* Naturally, when we compress the material on both sides, we expect the material to bulge out in the other two axes. In other words, we expect  $\epsilon_{22}$  and  $\epsilon_{33}$  to be nonzero. If  $S$  is a type (0, 0) tensor, this means that

$$\epsilon_{ij} = S\sigma_{ij}$$

meaning that  $\epsilon_{11}$  is the only nonzero component. But this cannot be the case, since we expect  $\epsilon_{22}$  to be nonzero.

If  $S$  is a type (1, 1) tensor, this means that:

$$\epsilon_{ij} = S_i^l \sigma_{lj}$$

If we compute  $\epsilon_{22}$  from this:

$$\epsilon_{22} = S_2^l \sigma_{l2}$$

and since only  $\sigma_{11} \neq 0$ , then we're forced to also conclude that  $\epsilon_{22} = 0$ . Again, this is impossible, so  $S$  cannot be a type (1, 1) tensor either. As for the nonzero entries, we see that all entries which have  $j = 1$  can be nonzero, which is to say that the first column of  $\epsilon$  are the only entries that can be nonzero.

Finally, if  $S$  is a type (2, 2) tensor:

$$\epsilon_{ij} = S_{ij}^{lm} \sigma_{lm}$$

Here, the value of  $\epsilon_{ij}$  is determined entirely by our tensor  $S$ , since for every entry of  $\epsilon$  we are summing over the entirety of  $\sigma$ . This is good, since it means that  $\epsilon_{22}$  can be nonzero. In fact, since  $S$  can be chosen freely, all the components of  $\epsilon$  can be nonzero.  $\square$

---

You should have found that the stiffness tensor must be a type-(2,2) tensor. Due to the form of Hooke's law, the symmetry properties of the stress and strain tensors get inherited by the stiffness tensor. In particular,

$$\begin{aligned} C_{k\ell}^{ji} &= C_{k\ell}^{ij} && \text{Symmetry in the two upper indices.} \\ C_{\ell k}^{ij} &= C_{k\ell}^{ij} && \text{Symmetry in the two lower indices.} \end{aligned}$$

(e) How many total components will a type-(2,2) tensor in 3D space have? Given the symmetry in the two upper indices, show or argue that there are only six independent choices for the pair of indices  $(ij)$  and similarly for  $(k\ell)$ . How many independent components does this leave the stiffness tensor? [Note: For example, the pair  $(ij) = (12)$  and  $(ij) = (21)$  are not independent since  $C^{12}_{k\ell} = C^{21}_{k\ell}$ .]  
[Spoilers! The six components are  $(11)$ ,  $(22)$ ,  $(33)$ ,  $(23)$ ,  $(13)$ ,  $(12)$ .]

*Solution:* In 2D space, we have 4 indices and 3 values to choose from for each index, so we have  $3^4 = 81$  total possible components. Now given the symmetry of the tensor, we find that the components  $(i, j) = (k, \ell) = (1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)$  are the only independent components, since all the other combinations of  $i$  and  $j$  are symmetric to one of these six pairs. Therefore, for each pair of indices, there are only 6 independent parameters.

In total, this leaves the stiffness tensor with  $6 \times 6 = 36$  independent parameters. □

---

The six possible pairs for the upper pair of indices  $(ij)$  can act as a single row index and the six possible pairs for the lower pair of indices  $(k\ell)$  can act as a single column index, allowing us to express the stiffness tensor as a  $(6 \times 6)$  matrix. We can similarly express the six independent components of the strain and stiffness tensors as six-component row vectors. For example,  $\sigma \doteq (\sigma_{11} \ \sigma_{22} \ \sigma_{33} \ \sigma_{23} \ \sigma_{13} \ \sigma_{12})$ . This is known as the ***Voigt form*** for stress, strain, and stiffness.<sup>9</sup> There is an additional symmetry that becomes apparent when we express the stiffness tensor in Voigt form which is that the  $(6 \times 6)$ -stiffness matrix  $\mathbf{C}^{(\text{mat})}$  is a *symmetric* matrix.

(f) Given this new symmetry, how many total independent components remain for the stiffness tensor?

*Solution:* Written in this form, we essentially have a symmetric,  $6 \times 6$  matrix, meaning that the total number of independent components is given by the formula for the number of free parameters in a symmetric matrix of dimension  $n$ :

$$N = \frac{n(n+1)}{2}$$

Since this matrix has 6 rows and columns, we sub  $n = 6$ :

$$N = \frac{6 \times 7}{2} = 21$$

this formula is a general formula to find the number of independent parameters for any symmetric matrix, so we plug in the dimension of 6 to get our answer. □

---

***Isotropic*** materials are characterized by properties which are independent of direction. Glass is an isotropic material, for example, but something like wood, which has a defined grain direction, is anisotropic. Isotropy and the symmetries of the stiffness tensor imply that stiffness must be a linear combination of  $g^{ij}g_{k\ell}$  and  $\delta_k^i\delta_\ell^j + \delta_\ell^i\delta_k^j$ . Thus, only *two* free scalar parameters remain in the stiffness tensor:

- $K$ , the ***bulk modulus***,
- $\mu$ , the ***shear modulus***.

In terms of these parameters, the components of the stiffness tensor are

$$C^{ij}_{k\ell} = (K - \frac{2}{3}\mu) g^{ij}g_{k\ell} + \mu(\delta_k^i\delta_\ell^j + \delta_\ell^i\delta_k^j) \quad (3)$$

---

<sup>9</sup>The compliance tensor in this case can also be written as a  $(6 \times 6)$ -matrix and can be found as the matrix inverse of the  $(6 \times 6)$  stiffness matrix. See [https://en.wikipedia.org/wiki/Voigt\\_notation](https://en.wikipedia.org/wiki/Voigt_notation) for more.

(g) An isotropic material undergoes a stress and experiences a strain  $\epsilon_{12} = \epsilon_{21} = a$ ,  $\epsilon_{22} = b$  (all other components are zero). Use Hooke's law and Eq. 3 to determine the components of the stress tensor  $\sigma_{ij}$ .

*Solution:* Hooke's law states:

$$\sigma_{ij} = C_{ij}^{lm} \epsilon_{lm}$$

So expanding  $C$  in terms of equation 3, we have:

$$\begin{aligned} \sigma_{ij} &= \left[ \left( K - \frac{2}{3}\mu \right) g^{lm} g_{ij} + \mu (\delta_i^l \delta_j^m + \delta_j^l \delta_i^m) \right] \epsilon_{lm} \\ &= \left( K - \frac{2}{3}\mu \right) g^{lm} g_{ij} \epsilon_{lm} + \mu \delta_i^l \delta_j^m \epsilon_{lm} + \mu \delta_j^l \delta_i^m \epsilon_{lm} \\ &= \left( K - \frac{2}{3}\mu \right) g^{lm} g_{ij} \epsilon_{lm} + \mu (\epsilon_{ij} + \epsilon_{ji}) \end{aligned}$$

And since we're working in cartesian space, the metric is the Kronecker delta:

$$\sigma_{ij} = \left( K - \frac{2}{3}\mu \right) \delta^{lm} \delta_{ij} \epsilon_{lm} + \mu (\epsilon_{ij} + \epsilon_{ji})$$

Now we can start plugging in the indices for  $i, j$ . Note that since we know  $\sigma$  is a symmetric matrix, we only have to plug in  $(i, j) = (1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)$ . Plugging this in:

$$\begin{aligned} \sigma_{11} &= \left( K - \frac{2}{3}\mu \right) \delta^{lm} \epsilon_{lm} + \mu (\epsilon_{11} + \epsilon_{11}) = \left( K - \frac{2}{3}\mu \right) \epsilon_{22} = \left( K - \frac{2}{3}\mu \right) b \\ \sigma_{12} &= \left( K - \frac{2}{3}\mu \right) \delta^{lm} \underbrace{\delta_{12}}_{=0} \epsilon_{lm} + \mu (\epsilon_{12} + \epsilon_{21}) = 2\mu \epsilon_{12} = 2a\mu = \sigma_{12} \\ \sigma_{13} &= \left( K - \frac{2}{3}\mu \right) \delta^{lm} \underbrace{\delta_{13}}_{=0} + \mu (\underbrace{\epsilon_{13} + \epsilon_{31}}_{=0}) = 0 + 0 = 0 = \sigma_{31} \\ \sigma_{22} &= \left( K - \frac{2}{3}\mu \right) \delta^{lm} \epsilon_{lm} + \mu (\epsilon_{22} + \epsilon_{22}) = \left( K - \frac{2}{3}\mu \right) \epsilon_{22} + 2\mu \epsilon_{22} = b \left( K + \frac{4}{3}\mu \right) \\ \sigma_{23} &= \left( K - \frac{2}{3}\mu \right) \delta^{lm} \underbrace{\delta_{23}}_{=0} \epsilon_{lm} + \mu (\underbrace{\epsilon_{23} + \epsilon_{32}}_{=0}) = 0 + 0 = 0 = \sigma_{32} \\ \sigma_{33} &= \left( K - \frac{2}{3}\mu \right) g^{lm} \epsilon_{lm} = b \left( K - \frac{2}{3}\mu \right) \end{aligned}$$

The other components of  $\sigma$  are automatically determined from this, due to the symmetry. □

---



---



## 7.5 - Spherical Coordinates

*Relevant Videos: The Metric; The Jacobian and Coordinate Changes*

When we have a tensor field (a tensor that depends on where you are in space), we can do a coordinate change from old coordinates  $\{x^i\}$  to new coordinates  $\{\tilde{x}^i\}$  by defining the **Jacobian matrix** and the **inverse Jacobian matrix**,

$$J^i_j \equiv \frac{\partial x^i}{\partial \tilde{x}^j}, \quad (J^{-1})^i_j \equiv \frac{\partial \tilde{x}^i}{\partial x^j}.$$

Upon a change-of-coordinates, contravariant indices (upper indices) transform via  $J^{-1}$  and covariant indices (lower indices) transform via  $J$ . Consider the transformation from Cartesian coordinates  $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$  to spherical polar coordinates  $\begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix}$ . Remember that the *physicist's* way of defining spherical coordinates has  $\theta$  as the polar angle and  $\varphi$  as the azimuthal angle,

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta. \end{aligned}$$

In Cartesian coordinates the components of the metric tensor are  $g_{ij} = \delta_{ij}$ .

(a) Find the Jacobian matrix  $J$ .

[*Supplementary Part (Not for Credit): Compute the inverse of the Jacobian  $J^{-1}$ .*]

*Solution:* Taking the derivatives, we get:

$$\begin{array}{lll} J_1^1 = \sin \theta \cos \phi & J_2^1 = r \cos \theta \cos \phi & J_3^1 = -r \sin \theta \sin \phi \\ J_1^2 = \sin \theta \sin \phi & J_2^2 = r \cos \theta \sin \phi & J_3^2 = r \sin \theta \cos \phi \\ J_1^3 = \cos \theta & J_2^3 = -r \sin \theta & J_3^3 = 0 \end{array}$$

Arranging this into a matrix, we get:

$$J = \begin{pmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

□

(b) Use the Jacobian matrix to find the components of the metric tensor in spherical coordinates. You may put the components of the type-(0,2) metric tensor in matrix form,  $g^{(\text{mat})} = \begin{pmatrix} g_{rr} & g_{r\theta} & g_{r\varphi} \\ g_{\theta r} & g_{\theta\theta} & g_{\theta\varphi} \\ g_{\varphi r} & g_{\varphi\theta} & g_{\varphi\varphi} \end{pmatrix}$ .

*Solution:* We have the relation that  $g_{\text{new}}^{\text{mat}} = J^\top g_{\text{old}}^{\text{mat}} J$ . Since the old basis was Cartesian, we know that  $g_{\text{old}}^{\text{mat}} = \delta_{ij} = \mathbb{1}$ . Therefore, we find that  $g_{\text{new}}^{\text{mat}} = J^\top J$ . Computing this matrix multiplication using a computer, we get:

$$g_{\text{new}}^{\text{mat}} = J^\top J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}$$

□

---

(c) Given a vector  $\vec{v} = v^r \hat{r} + v^\theta \hat{\theta} + v^\varphi \hat{\varphi}$ , find the components  $\tilde{v}_i \equiv g_{ij} v^j$  of the dual vector  $\underline{v}$  in spherical coordinates.

*Solution:* Just working through it by plugging in the indices:

$$\begin{aligned}\tilde{v}_1 &= g_{1j} v^j = v^r \\ \tilde{v}_2 &= g_{2j} v^j = r^2 v^\theta \\ \tilde{v}_3 &= g_{3j} v^j = r^2 \sin^2 \theta v^\varphi\end{aligned}$$

and we're done!

□

---

*Commentary:* The above part shows that in other coordinate systems or with non-trivial metrics, we can't just "transpose" our vector to find a dual vector!

◆