All About Tensors

Michael Astwood

Physics Club Colloquium, October 1st 2019

Tensors are a topic that scare a lot of students. I'm here to clear up that anxiety and clarify how tensors really aren't as bad as people think they are. To begin let's review a little linear algebra.

Definition 1 (Field) *A field is a special kind of set. We will denote a general field $\mathbb{K}$. What makes a field special is that you can add, subtract, multiply, and divide.*

1. *There is a multiplicative identity. That is, there is some number 1 so that $1a = a1 = a$.*
2. *You can divide by anything except zero. That is:$\forall a \neq 0 \in \mathbb{K}, \exists a^{-1}$ s.t. $a^{-1}a = aa^{-1} = 1$*
3. *There is an additive identity. That is: there is some number 0 so that $0 + a = a + 0 = a$*
4. *There are negative numbers. That is: for any a there is some $-a$ so that $a + (-a) = 0$.*
5. *Multiplication is distributive. That is: $a(b + c) = ab + ac$.*
6. *Addition and multiplicative are associative. $a(bc) = (ab)c$, $a + (b + c) = (a + b) + c$*
7. *Addition is commutative. $a + b = b + a$*

Some examples of fields include $\mathbb{R}, \mathbb{C}$, or $\mathbb{Q}$, but there are more wacky fields (such as finite fields) which have strange properties.

Definition 2 (Vector Space) *Recall the definition of a vector space V over a field $\mathbb{K}$. A vector space is a set with addition (with all its' usual properties) where we have introduced a concept of "scalar multiplication".*[1]

1. *The space is linearly closed: $a\vec{x} + b\vec{y} \in V$*
2. *The zero scalar coincides with the zero vector: $0\vec{a} = \vec{0}$.*
3. *The identity scalar is compatible with V: $1\vec{v} = \vec{v}$*

Now that we have a vector space, it's natural to think about maps between vector spaces. Everyone here knows about linear operators, which take V to V. A tensor is a convenient way of defining a linear map which takes any number of inputs. Lets' start with the simplest case of a linear map, and we will see in a few minutes that there is a way to build any tensor out of so-called tensor products of these maps.

Definition 3 (Dual Space) *Consider a linear map $\alpha : V \to \mathbb{K}$. Such a linear map has various names (which I will list in the footnotes*[2]*). Linear maps of this type form a natural vector space called the Dual Space to V.*

$$V^* \equiv \{\alpha : V \to \mathbb{K}\} = L(V, \mathbb{K})$$

$$(a\alpha + b\beta)(\vec{v}) = a\alpha(\vec{v}) + b\alpha(\vec{v})$$

[1] In the definition of a vector space, if we replace a field with a ring, we get something called an R-module (where R is the ring).

[2] Covector, Row Vector, Ket, Linear Functional, (0,1) Tensor, etc

It's quite simple to build a basis for the Dual Space. Simply take a basis $\mathcal{B} = \{e_1, ..., e_n\}$ for V and consider the maps $e^1, ..., e^n$ so that $e^i(e_j) = \delta_i^j$.

$$e^i(e_j) = \delta_j^i \equiv \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Of course, we can not guarantee that there is only one such choice of e_i unless we define a few more things. We will get to that later. However, one can already prove that $\dim(V^*) = \dim(V)$. One can write an arbitrary covector as $\alpha = \sum \alpha_i e^i$.

Now we are ready to define a Tensor. First we'll give a broad definition, and then see how all of the fancy physics applications follow from the definition.

Definition 4 (Tensor) *An (m,n)-tensor is a multilinear map $T : V \times ... \times V \times V^* \times ... \times V^* \to \mathbb{K}$, where V is repeated n times and V^* is repeated m times. Multilinear means that the map is linear in each argument:*

$$T(v_1, ..., av_i + bw_i, ..., v_n, \alpha_1, ..., \alpha_m) = aT(v_1, ..., v_i, ..., v_n, \alpha_1, ..., \alpha_m) + bT(v_1, ..., w_i, ..., v_n, \alpha_1, ..., \alpha_m)$$

Definition 5 (Tensor Space) *The set of (m,n) tensors is denoted by the Tensor Product of vector spaces. That is:*

$$V \otimes ... \otimes V \otimes V^* \otimes ... \otimes V^* = \{T : (V^*)^n \times (V)^m \to \mathbb{K}\}$$

One of the most important features of a tensor is the basis from which it is constructed. Let's take as a simple example the $(1,1)$-tensors (which will be familiar to you in a second). We define a tensor product of vectors e_i, e^j as follows.

$$(e^i \otimes e_j)(a, \alpha) = e^i(a) e_j(\alpha) \equiv e^i(a) \alpha(e^j)$$

Let's examine this in $\mathbb{R}^2$. This gives a very familiar result from PHYS234. Here's an example:

$$e^1 \otimes e_1 = (e_1)^T e_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$(e_1 \otimes e^1)(a, \alpha) = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Thus a basis for $V^{\otimes n} \otimes (V^*)^{\otimes m}$ is given by $\mathcal{B}_{\otimes} = \{e_{j_1} \otimes ... \otimes e_{j_n} \otimes e^{i_1} \otimes ... \otimes e^{i_m}\}$. There is a natural way to define the way a tensor acts on another tensor. We will define this next.

Definition 6 (Tensor Product of Maps) *Let L be a linear map from V to W, and let M be a linear map from A to B. We define their tensor product to be:*

$$L \otimes M : V \otimes W \to A \otimes B$$

$$(L \otimes M)(v \otimes w) = L(v) \otimes M(w)$$

This provides an interesting result. $e^i \otimes e^j : V \otimes V \to \mathbb{K}$ is the same as saying $e_i \otimes e_j : V \times V \to \mathbb{K}$. Therefore we have a natural correspondence (which you might have thought was obvious, but only follows from the previous definition!)

$$V \otimes V \cong (V^* \otimes V^*)^*$$

A great example of the use of these in physics is in quantum mechanics. The first example is one you saw in PHYS234 (if you've taken it already). If you haven't seen braket notation, now is the absolute best time to learn it:

$$|\psi\rangle \in \mathcal{H} \iff \langle\psi| \in \mathcal{H}^*$$

$$|\psi\rangle\langle\phi| \equiv |\psi\rangle \otimes \langle\phi|$$

$$|\psi, \phi\rangle = |\psi\rangle |\phi\rangle \equiv |\psi\rangle \otimes |\phi\rangle$$

From this, one can see that the braket notation is a great way to condense tensor product notation so that one no longer needs to spend hours typing "otimes" repeatedly in latex. The two most common ways to write tensors as a physicist will be with braket notation and index notation which we will see in a second.

Definition 7 (Entanglement) *A quantum mechanical system with multiple non-interacting subsystems is described as the tensor product of the subsystems: $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$. An entangled state is a state $|\psi\rangle$ which can not be decomposed into the tensor product of two states $|a\rangle, |b\rangle$. For example, the following two-particle spin state is entangled:*

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|-\rangle_A |+\rangle_B - |+\rangle_A |-\rangle_B)$$

All of these calculations we've seen can be captured quite easily through the lens of Einstein's index notation. We define the following notation so that we don't have to write a billion Σs.

$$a_i b^i \equiv \sum_{i \in I} a_i b^i$$

Whenever we work with index notation, we will use upper indices for vectors in V, and lower indices for vectors in V^*. This is just convention.

Now we can write tensors in index notation:

$$A \equiv \sum_{i,j} A^i_j e_i \otimes e^j \implies \alpha A\vec{v} = \alpha_j A^j_i a^i = A^j_i \alpha_j a^i$$

As such, high-ranking tensors can be written in a much more compact form.

$$A \equiv A^{j_1 \ldots j_m}_{i_1 \ldots i_n} e_{j_1} \otimes \ldots \otimes e_{j_m} \otimes e^{i_1} \otimes \ldots \otimes e^{i_n}$$

Index notation gives a convenient way to study a change of basis, as well as how we can convert vectors into covectors and vice versa. It also tells us how to use tensors in the way that you're used to - as linear maps between vector spaces rather than as functionals. I'm not going to cover change of basis in detail, since you will need to know something about manifolds to get the full picture in terms of derivatives and so on. However, I will show you why covectors and vectors are distinguished even in flat space.

Here's how a tensor acts on a single input. Suppose you have an (n, m) tensor (with $n \neq 0$) T and a vector v.

$$Tv \equiv \sum_{i_k} T^{j_1, \ldots, j_m}_{i_1, \ldots, i_n} v^{i_k} = (Tv)^{j_1, \ldots, j_m}_{i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n}$$

So we see that acting a tensor on another tensor effectively has the effect of producing a new tensor with smaller rank - this is what we mean when we say that indices contract. Now let's see what these indices mean physically.

Suppose we change our coordinates so that $e'_j = A^j_i e_i$.

$$v^i e_i = v'^i e'_i \implies v^i e_i = v'^i A^j_i e_i$$

So $v'^i = (A^{-1})^j_i v^j$. As such, we say that vectors transform contravariantly - when you change to a new coordinate system you must introduce the inverse of the transformation matrix to change the components. Similarly, for covectors we have:

$$\alpha(v^i e_i) = \alpha(v'^i e'_i) \implies \alpha_i v'^i A^i_j e^i(e_i) = \alpha_i v^i$$

So $A^i_j \alpha'_i = \alpha_i$, which is what we call covariant. This is why we refer to covectors as covectors - there is a clear distinction between how these transform versus regular vectors. This should be simple enough mathematically to extend to tensors - simply break up a tensor into a sum of tensor products and apply the transformation to each component. I will leave this as an exercise for the sake of time, but note how a tensor is made up of contravariant and covariant components multiplied (then summed) together. This will tell you something about how to break up the product.

Now let's talk about raising and lowering indices. First we need a tool to do this, called the metric.

Definition 8 (Metric Tensor) *A pseudo-Riemannian metric is a symmetric tensor $g_{ij} = g_{ji} = (g^{-1})^{ij}$ with the property that $g_{ij} v^i w^j = 0$ only when v or w is the zero vector.*

If we define a metric on our space, we immediately have a method to create covectors from vectors and vice versa. This provides a basis-free isomorphism between V and V^*. The way we do this is by "contracting indices": $g_{ij} v^i \equiv v_j$, so that $v_j e^j \equiv v_j g(e_j, \cdot)$. Studying metrics is a very rich field in differential geometry, where the transformation rules become quite a bit more complicated due to the curvature of the space you're living in.

Another topic we can cover is the very important decomposition of a tensor into its' symmetric and antisymmetric components. Let's look at 2-tensors as a first example.

$$T \equiv T_{ij} e^i \otimes e^j = \frac{T_{ij}}{2}(e^i \otimes e^j + e^j \otimes e^i) + \frac{T_{ij}}{2}(e^i \otimes e^j - e^j \otimes e^i) \equiv \frac{1}{2}(T_{[ij]} + T_{\{ij\}})$$

The first component is called the symmetric part of the tensor, and the second component is called the antisymmetric part. We can (anti)symmetrize any size of tensor (in a roundabout way) by doing a horrible looking sum over permutations. You use this notation for it:

$$T_{[ijk]l} = T_{[ikj]l} = T_{[kij]l} = ...$$

We won't do any big examples because they can get arbitrarily complicated. Many important quantities in physics are computed by taking symmetric and antisymmetric components of tensor products. We have an important name for the second component, however, called the wedge (or exterior) product.

$$e^i \wedge e^j \equiv e^i \otimes e^j - e^j \otimes e^i$$

The space of antisymmetric tensors is called the Grassman/Exterior Algebra, and is defined as follows.

Definition 9 (Grassman Algebra) *Let $\Lambda^0(V)$ be the field over which V is defined. Similarly, let $\Lambda^1(V) = V$. We define the Grassman algebra as the following direct sum of vector spaces, as an algebra over $\mathbb{K}$ given by the wedge product. Elements of $\Lambda^k(V^*)$ are called k-forms, and elements of $\Lambda^k(V)$ are called k-blades (or k-vectors).*

$$\Lambda^n(V) \equiv \left\{ \sum a_{i_1,\dots,i_n} \vec{v}_1 \wedge \dots \wedge \vec{v}_n : v_i \in V \right\}$$

$$\Lambda^\bullet(V) \equiv \bigoplus_{n=0}^{\infty} \Lambda^n(V)$$

This algebra is one method to derive a rigorous explanation for the existence/nonexistence of the cross product in certain vector spaces. It's also the foundation for the theory of differential forms, which are forms over the vector space of differential operators on a manifold, with coefficients in $C^n(M)$. With these we can define contracted versions of the Maxwell equations as well as their advanced equivalents - the Yang-Mills equations. Not to mention all the GR you can do with this stuff.

Another useful object is the interior product of a vector with a tensor. This comes up in a bunch of interesting formulas and it has a bunch of wacky properties, but it's not something you see too too often.

Definition 10 (Interior Product) *We define the interior product $\iota_v \alpha$ of v with a tensor α to be the following tensor:*

$$(\iota_v \alpha)(a_1, ..., a_n) = \alpha(v, a_1, ..., a_n)$$