*Notes for GR course (Rafael D. Sorkin, 2013)*

## WHAT ARE TENSORS?

One can understand a tensor as being *a sum of products of vectors and covectors*. This is perhaps the concept's most basic meaning, but one can also think of tensors in many other ways, this multiplicity being one source of their usefulness. To list some of the most important ways, you can think of a tensor as:

• A sum of products of vectors and covectors

• A linear operator from one space of tensors to another

• A multilinear form taking a number of vectors and covectors to a scalar

• A certain kind of infinitesimal geometrical object

• An array of numbers that transforms in a certain way under change of basis

• An element of a vector space that arises as the solution of a certain "universal mapping problem" involving multilinear mappings of vectors and covectors.

• A symbol carrying some indices

In these notes we will concentrate on the first and second meanings. The third is found in many textbooks (including Bob Wald's). The fourth generalizes the ideas of a vector as an infinitesimal displacement (a pair of points) and of a covector as an infinitesimal "parallel plate capacitor". The fifth (is it the most familiar to you?) simply defines a tensor in terms of its components relative to a basis; it generalizes the representation of a linear operator as a matrix. The sixth comes from category theory. The seventh is partly a joke, but not wholly.

Notice that a metric regarded as a quadratic form is a special case of the third meaning.

When we say that a tensor is a product of vectors it is not necessary that all the vectors come from the same space. However, I will limit myself here to vectors taken from the spaces $V$ and $V^*$, where $V$ will usually be the tangent space to a manifold and $V^*$ is its dual space. (That is $V^* = L(V, \mathbb{R})$, the space of linear mappings taking elements of $V$ into scalars. Recall that an element of $V^*$ is called a *covector*.) For us, $V$ will always have finite dimension.

Let's start with the simplest space of tensors we can make in this way, namely $V \otimes V$, the space of sums of products of pairs of vectors. The most general element of $V \otimes V$ can be written as a finite sum

$$T = \sum_n u[n] \otimes v[n] \tag{1}$$

where $u[n]$ and $v[n]$ are vectors in $V$. (One might think, momentarily, that a more general possibility would be $T = \sum_n \lambda[n]\, u[n] \otimes v[n]$, $\lambda[n]$ being a scalar; but, clearly, this reduces to (1) because we can absorb $\lambda[n]$ into either $u[n]$ or $v[n]$.) (Notice that the same symbol $\otimes$ is used to denote both products of whole spaces and products of their elements.)

But what kind of product, exactly, is '$\otimes$'? It is not the scalar product of course (which anyhow is not even defined between $V$ and $V$ in the absence of a metric). Rather it is just something we invent to allow ourselves to multiply vectors, without imposing any other conditions than the rules of distributivity and associativity that are implicit in the idea of product itself. (It is called sometimes "tensor product" sometimes "outer product".) In other words we have, $u \otimes (v + w) = u \otimes v + u \otimes w$, $u \otimes (v \otimes w) = (u \otimes v) \otimes w$, $(\lambda u) \otimes v = \lambda (u \otimes v)$ ($\lambda$ being a scalar), $u \otimes 0 = 0$ (0 being the zero tensor), etc. These rules let us determine when two different expressions of the form (1) produce the same tensor $T$, and that is all we need to give our sum of products a precise meaning.

So how can we decide whether two sums $T_1$ and $T_2$ of the form (1) are equal? Clearly, they must be equal if we can convert one into the other by a finite number of rearrangements using distributivity, associativity, etc. The idea, then, is to make this also a necessary condition for equality: we declare that if we cannot convert $T_1$ into $T_2$ this way, then they are *not* equal. In practice it is often easy to decide the question by just playing around, but for definiteness, it's nice to have a mechanical procedure that always works. Here is a simple algorithm that does. (The generalization from twofold products to multiple products should be obvious.)

1. Let $T = T_1 - T_2$ . Clearly $T_1 = T_2 \iff T = 0$ .

2. If the factors $u[n]$ in $T$ are not all linearly independent, then pick a subset of them which is, express the remaining $u$'s in terms of these, and expand out the resulting expression into a sum of simple products using distributivity. This yields an expression which is again in the form (1), but now all the surviving $u$'s are linearly independent.

3. If any nonzero terms remain, then $T \neq 0$, otherwise of course $T = 0$ (and consequently $T_1 = T_2$).

Sometimes, when one is trying to prove some general theorem of tensor algebra (e.g. that contraction is well-defined in general) it is useful to have a more compact form of the criterion for $T = 0$ implicitly defined by the above algorithm. It is not too hard to see that the following is such a criterion: $T = 0$ iff there exists a matrix $M[mn]$ such that $\sum_n M[mn]\, v[n] = v[m]$ and $\sum_m u[m]\, M[mn] = 0$.

Now that we have given precise meaning to tensors, defining their algebra is pretty trivial, with one exception. We can add tensors, take their outer products, multiply them by scalars, all without difficulty. The partial exception is the process of *contraction*, which reduces the number of factors in a tensor product by combining factors from $V$ and $V^*$ to form scalars.

Before defining contraction, however, it will be useful to say something about the hardest topic in tensor algebra: the notation! We will be using the so called "abstract index notation", which decorates a symbol like $T$ with a certain number of *contravariant* indices (superscripts) and *covariant* indices (subscripts) that indicate how the tensor was built up from $V$ and $V^*$. Specifically, each contravariant index corresponds to a factor from $V$ (a vector) and each covariant index to a factor from $V^*$ (a covector). Thus, for example, a symbol like $T^{ab}{}_c$ denotes a tensor in $V \otimes V \otimes V^*$, and $g_{ab}$ a tensor in $V^* \otimes V^*$. With this notation, the symbol $\otimes$ becomes superfluous in most situations, and we can write the outer product using juxtaposition, just as we ordinarily do for ordinary multiplication. Thus, for example, we can write $u^a v^b$ instead of $u \otimes v$. Notice that a vector is now a tensor with one contravariant index and a covector is a tensor with one covariant index. A scalar, is of course, the same thing as a tensor with no indices at all.

*Contraction* is now, symbolically, a process in which we cancel one or more contravariant indices against an equal number of covariant indices with which they are matched. For example, from a vector $v^a$ and a tensor $g_{ab}$ we can form a covector $u_a = g_{ab} v^b$ by "contracting" the second index of $g$ with the first (and only) index of $v$. Notice that we indicate which indices are matched for contraction by writing them with the same letter. The mathematical meaning of this operation can be built up inductively from the simplest case of contraction of a covector with a vector, which in turn has the obvious definition,

$$\omega_a v^a = (\omega, v) \ , \qquad (2)$$

the result of applying the linear function $\omega$ to the vector $v$.

Actually, since $V$ is finite dimensional, we know that $V = (V^*)^*$, so that one can equally well think of covectors as basic and vectors as derived from them by dualization. For this reason it makes sense to write the *inner product* (as it's called) between a vector and a covector in the more neutral way done in (2), rather than as $\omega(v)$. (It's lucky that $V^{**} = V$. If they were different, we'd need a minimum of three types of indices to represent the most general tensor.)

It's now clear how contraction is to be defined in general. We just write out every tensor as a sum of products of vectors and covectors, and, whenever we have to contract a pair of indices, we take the inner product of the corresponding vector and covector in each term of the sum. For example, if $T^a{}_b = u^a v_b - v^a u_b$ (a so called wedge product), then

$$T^a{}_b w^b = u^a (v_b w^b) - v^a (u_b w^b) \ .$$

If you are mathematically scrupulous, you might worry whether this definition is compatible with that of equality of tensors: might changing the way we represent a tensor as a sum of products change the result of contracting two of its indices? Fortunately it is easy to see that this is not the case, using, for example, the simple criterion for equality stated above. By the way, an index which is contracted is called a *dummy index*, otherwise it is a *free index*. Thus in the above equation, $a$ is free and $b$ is a dummy.

Given these basic definitions, the rest of the story is pretty straightforward.

The abstract index notation (in conjunction with the idea of contraction of indices) exhibits clearly how a tensor can be interpreted (in many different ways) as a linear operator from one tensor space to another. We saw above how a covariant tensor $g_{ab}$ yields a mapping from $V$ to $V^*$ (in fact two different such mappings if $g$ is not symmetric). It similarly yields a map from $V \otimes V$ to scalars, and for that matter, a map from scalars to $V^* \otimes V^*$ (namely the map $\lambda \to \lambda g_{ab}$ !). Notice, here, that the expression $g_{ab} u^a v^b$ can be read in at least two ways: either as the separate contraction of two vectors with a tensor to yield a scalar (this is then the bilinear "scalar product" of the two vectors with respect to $g$) or as a contraction of the tensor $u^a v^b$ with the tensor $g_{ab}$. As a final example consider a mixed tensor $T^a{}_b$. It yields, among other things, linear maps from $V$ to $V$ and from $V^*$ to $V^*$. All this shows how tensors may be interpreted as linear operators. The converse, that every linear operator from one tensor space to another arises in this way, follows easily from dimension counting, using the fact that

$$\dim(V \otimes W) = \dim(V) \dim(W) \ .$$

(This in turn follows by noting that the outer products of basis vectors for $V$ and $W$ give us a basis for $V \otimes W$.)

So far, we have not mentioned components of tensors, and indeed, we cannot even speak about them before choosing a basis for $V$, so imagine that we have done this. *Relative to such a basis*, we can define the *components* of any tensor $T$ by expanding it in terms of tensor products of basis vectors and covectors. (Equivalently, we can obtain the components by contracting $T$ in all possible ways with basis covectors and vectors.) A one-index tensor then appears as a column or row vector, a two-index tensor as a matrix, etc. The indices on a tensor can now be interpreted "concretely" as numbers labeling the components. Instead of being just decorations they now run from 0 to 3 (or whatever). Tensor product becomes just multiplication of the components, and an equation like $w^{ab} = u^a v^b$ becomes just a set of numerical equations, one for each pair of values for $a$ and $b$.

Most everything having to do with components follows easily from the definition of the dual basis together with the familiar "completeness relation", which lets us insert factors of unity in convenient locations. The formulas for how tensors transform when we change from one basis for $V$ to another are easily derived in this manner. (These formulas are the origin of the "contravariant/covariant" terminology.) Similarly, it is easy to prove that the contraction of two indices now amounts to setting them both equal to the same number $\mu$ and then summing over the possible values of $\mu$. Our rule for using equal letters to show which index is contracted with which other one thus becomes the familiar "summation convention".

Finally, a word about symmetry properties of tensors. A tensor like $v \otimes v$ is obviously symmetric under exchange of its factors. More generally given any sum of the form (1) we can obtain a new tensor $T'$ by exchanging the factors $u$ and $v$. (Exercise: prove that this operation on tensors is well defined.) In the abstract index notation, we have $T^{cd} = \sum_n u[n]^c \, v[n]^d$ and $T'^{cd} = \sum_n v[n]^c \, u[n]^d = \sum_n u[n]^d \, v[n]^c$. Thus it is natural to express the relationship between $T$ and $T'$ by writing $T'^{cd} = T^{dc}$. (In terms of components this just says that the matrices representing $T$ and $T'$ are transposes of each other.) More generally, we can permute the indices of any tensor to produce a new tensor (but the permutation must not mix contravariant indices with covariant ones). A tensor which transforms simply under such permutations is said to obey a symmetry condition. The simplest possibilities are total symmetry and total antisymmetry (skewness), but the Riemann tensor, for example, belongs to a more complicated symmetry type.