18.024–ESG Notes 3

Pramod N. Achar

Spring 2000

In this set of notes, we will develop the basic theory of tensors and differential forms; we will learn what it means to integrate a differential form; and we will state the generalized Stokes' Theorem in terms of differential forms. *Caveat lector*: there are several closely related meanings of the word *tensor*.

We begin introducing some convenient terminology for which we will not give precise definitions. A *manifold* is a curve, surface, or higher-dimensional generalization thereof. We will often shorten the phrase "manifold of dimension k" to "k-manifold." When we integrate over a k-manifold M, we will need to make use of a *parametrization* $\mathbf{r} : D \to M$, where $D \subset \mathbb{R}^k$. If $M \subset \mathbb{R}^n$, then we say the *codimension* of M is $n - k$. Thus, curves are 1-manifolds. 2-manifolds sitting in $\mathbb{R}^3$ have codimension 1, but if they sit in $\mathbb{R}^4$, they have codimension 2. Note that it makes sense to speak of "normal vectors" to manifolds of codimension 1.

In this set of notes, higher-dimensional integrals will never be written as multiple integrals.

Definition 1. A *tensor* of degree k, or a k-tensor, on $\mathbb{R}^n$ is a real-valued function of k variables, where each input variable is a vector in $\mathbb{R}^n$, such that the function is linear in each variable. That is, a tensor is a function

$$T : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k} \to \mathbb{R},$$

such that

$$T(\mathbf{v}_1, \ldots, c\mathbf{v}_i + c'\mathbf{v}'_i, \ldots, \mathbf{v}_k) = cT(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_k) + c'T(\mathbf{v}_1, \ldots, \mathbf{v}'_1, \ldots, \mathbf{v}_k)$$

for each i. By convention, a 0-tensor is a just a real constant.

The first thing to notice about tensors is that because they must be linear in each variable, then if we specify the values of a tensor on the unit coordinate vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$, then that determines the tensor uniquely.

What does this mean? Let us consider the simplest nontrivial case, that of a 1-tensor. This is just a linear function $T : \mathbb{R}^n \to \mathbb{R}$. If we know, for example, that $T(\mathbf{e}_1) = a_1, \ldots, T(\mathbf{e}_n) = a_n$, then we can compute $T(\mathbf{v})$ for any vector $\mathbf{v} \in \mathbb{R}^n$. Suppose that

$$\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n;$$

then, by linearity of T, we must have

$$T(\mathbf{v}) = v_1a_1 + \cdots + v_ma_n.$$

This expression can be written more concisely: if we let $\mathbf{a} \in \mathbb{R}^n$ be the vector $a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$, then

$$T(\mathbf{v}) = \mathbf{a} \cdot \mathbf{v}.$$

This discussion illustrates the relationship between vectors and tensors of degree 1. Indeed, the ideas in this paragraph could be made into a rigorous proof of the following fact.

Proposition 2. *There is a one-to-one correspondence between vectors in $\mathbb{R}^n$ and tensors of degree 1 on $\mathbb{R}^n$. The correspondence is as follows: if $T : \mathbb{R}^n \to \mathbb{R}$ is a tensor, the corresponding vector is*

$$(T(\mathbf{e}_1), \dots, T(\mathbf{e}_n));$$

conversely, given $\mathbf{a} \in \mathbb{R}^n$, the corresponding tensor is defined by

$$T(\mathbf{v}) = \mathbf{a} \cdot \mathbf{v}.$$

Some people, such as physicists, do not distinguish between a vector and a 1-tensor. We say that a tensor is determined by, or corresponds to, the list of n scalars $a_1, \dots, a_n$, whereas a physicist might just say that the tensor *is* the list of n scalars.

Let us now consider the next simplest case, that of tensors of degree 2. Again, because it is linear in each variable, the tensor $T : \mathbb{R}^n \times \mathbb{R}^n$ is determined by its values on the unit coordinate vectors. But because T is a function of two vector variables, we need to evaluate it on all pairs of unit coordinate vectors. Let

$$a_{ij} = T(\mathbf{e}_i, \mathbf{e}_j);$$

there are n^2 such numbers. As we might expect, 2-tensors are closely related to $n \times n$ matrices, but the algebra is more complicated than in the degree 1 case. The precise relationship is given in the following proposition, which we state without proof.

Proposition 3. *There is a one-to-one correspondence between $n \times n$ matrices and tensors of degree 2 on $\mathbb{R}^n$. The correspondence is as follows: if $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a tensor, the corresponding matrix is*

$$\begin{pmatrix} T(\mathbf{e}_1, \mathbf{e}_1) & T(\mathbf{e}_1, \mathbf{e}_2) & \cdots & T(\mathbf{e}_1, \mathbf{e}_n) \\ T(\mathbf{e}_2, \mathbf{e}_1) & T(\mathbf{e}_2, \mathbf{e}_2) & \cdots & T(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ T(\mathbf{e}_n, \mathbf{e}_1) & T(\mathbf{e}_n, \mathbf{e}_2) & \cdots & T(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix};$$

conversely, given an $n \times n$ matrix $\mathbf{A}$, the corresponding tensor is defined by

$$T(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot \mathbf{A}\mathbf{w}.$$

Again, physicists often think of the array of scalars a_{ij} as *being* the tensor, whereas mathematicians think of it as determining a tensor.

It is difficult to continue this analogy to tensors of degree greater than 2, but we can still make the observation that a tensor is determined by its values on the unit coordinate vectors. If T is a k-tensor on $\mathbb{R}^n$, consider the constants

$$a_{i_1 i_2 \dots i_k} = T(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_k}).$$

There are n choices for which unit coordinate vector to use in the first variable, n choices for the second, and so on, for a total of n^k constants $a_{i_1 i_2 \dots i_k}$. These n^k constants determine the tensor. It is possible to imagine these n^k constants arranged in a k-dimensional array with n rows in each direction, but when $k > 2$, this is not necessarily a useful way to think of tensors.

Given any two tensors S and T, there is an easy way to form a new tensor of higher degree out of them.

Definition 4. Let S be a k-tensor on $\mathbb{R}^n$, and T an l-tensor. Their *tensor product* $S \otimes T$ is a new tensor of degree $k + l$, defined by

$$(S \otimes T)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}) = S(\mathbf{v}_1, \dots, \mathbf{v}_k) T(\mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+l}).$$

Definition 5. An *alternating tensor* of degree k on $\mathbb{R}^n$ is a k-tensor with the property that exchanging any two of its input variables negates its value. That is, T is alternating if

$$T(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_k) = -T(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k).$$

This condition is vacuously satisfied by all tensors of degree 0 and 1, so we say that all degree 0 or 1 tensors are alternating. The set of all alternating k-tensors on $\mathbb{R}^n$ is denoted by $\bigwedge^k(\mathbb{R}^n)^*$.

It turns out that in the correspondence between 2-tensors and matrices, alternating tensors correspond precisely to skew-symmetric matrices. (This is why, in Problem Set 12, you integrated a skew-symmetric matrix field over a 2-manifold in $\mathbb{R}^4$—you were really integrating an alternating 2-tensor.)

We are now almost ready to give a general definition of integration.

Definition 6. Let $A \subset \mathbb{R}^n$. A *differential form* of degree k on A, also called a *k-form* or an *alternating k-tensor field*, is a function $\mathbf{f} : A \to \bigwedge^k(\mathbb{R}^n)^*$.

This last definition can be a little hard to get your head around: remember that the elements of $\bigwedge^k(\mathbb{R}^n)^*$ are tensors, which are themselves functions. So a differential form is a function that assigns to each point of its domain another function! This is a little less mind-boggling if we instead think of a form as assigning an array of n^k constants to each point, but it is important to remember that if ω is a k-form, and $\mathbf{x}$ is a point of its domain, then $\omega(\mathbf{x})$ is itself a function of k vector variables.

We take as given the definition and properties of k-dimensional integrals of scalar functions over regions of $\mathbb{R}^k$; *i.e.*, ordinary multiple integrals. We now define integration of differential forms over manifolds in terms of ordinary multiple integrals.

Definition 7. Let M be a k-manifold in $\mathbb{R}^n$, and let $\mathbf{r} : D \to M$ be a parametrization of it, where $D \subset \mathbb{R}^k$. Let ω be a k-form defined on some subset of $\mathbb{R}^n$ containing M. Let us write $t_1, \dots, t_k$ for the variables in $\mathbb{R}^k$. Then, we define

$$\int_M \omega = \int_D \omega\left(\frac{\partial \mathbf{r}}{\partial t_1}, \dots, \frac{\partial \mathbf{r}}{\partial t_k}\right).$$

As always, there is the question of whether this is well-defined; *i.e.*, whether the value of $\int_M \omega$ depends on the parametrization of M that we choose. This is why we need to require that ω be an *alternating* tensor field and not an arbitrary one. Here is the desired fact, without proof:

Theorem 8. *The integral of a k-form over a k-manifold is well-defined up to sign, independent of the parametrization.*

This fits well with what we have seen before for line integrals and surface integrals: a line integral changes sign if the curve is parametrized in the opposite direction, and a surface integral changes sign if the opposite normal direction is used. But in all these cases, the value of the integral is well-defined up to a sign.

The last major thing we need to do is to produce an analogue of the Second Fundamental Theorem. Suppose that M is a k-manifold with boundary N, a $(k-1)$-manifold; and let ω be a $(k-1)$-form. Following the pattern of the Second Fundamental Theorem, Green's Theorem, Stokes' Theorem, *etc.*, we expect a theorem that tells us that $\int_N \omega$ is equal to the integral of some related form over M, where that related form should be obtained by doing some sort of differentiation on ω. It should also be a form of degree one greater than that of ω, since M is a manifold of dimension one greater than that of N.

This is where we introduce the differential operator d. It has a uniform definition for all forms of all degrees in all dimensions, instead of idiosyncratic definitions like those we have seen for curl, divergence, *etc.* We will define it first for 0-forms, and then say how to compute the differential of a form of higher degree in terms of differentials of lower-degree forms. To achieve this reduction, we need a way to get higher-degree forms from lower-degree ones. The problem is that the tensor product operation $\otimes$ does not preserve the property of being alternating. That is, even if S and T are both alternating tensors, $S \otimes T$ almost never is.

Before we can fix this problem, we need to investigate in much greater detail what it means for a tensor to be alternating. An arbitrary k-tensor is determined by n^k constants, but most tensors are not alternating, so it should take a lot fewer constants to determine an alternating tensor. We have already seen this with 4×4 skew-symmetric matrices: it takes 16 numbers to determine a 2-tensor on $\mathbb{R}^4$, but only 6 numbers to determine an alternating 2-tensor on $\mathbb{R}^4$. (There are 6 numbers to specify in giving a 4×4 skew-symmetric matrix.)

How does this come about? To say T is alternating is to say that exchanging any two input variables changes the sign of the answer. It follows that if any two input variables to T are equal, the output value must be 0 (since it must equal its own negative). Thus, among the constants

$$a_{i_1 i_2 \dots i_k} = T(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_k}),$$

we get 0 whenever two of the indices $i_1, \dots, i_k$ are equal. (Hence the 0's along the diagonal of a skew-symmetric matrix.) The constants can be nonzero only if all k indices in the subscript are distinct. Moreover, the alternating property means that

$$a_{i_1 \dots i_p \dots i_q \dots i_k} = -a_{i_1 \dots i_q \dots i_p \dots i_k}.$$

Any rearrangement of the indices can be achieved by a series of exchanges like this, so there is really only parameter to be determined among all the possible rearrangements of $i_1 \dots i_k$. (Thus there 12 nonzero entries in a skew-symmetric matrix, but 6 of them are the negatives of the other 6, so there are really only 6 parameters to be determined.) Hence, the number of parameters needed to determine an alternating k-tensor is equal to the number of ways of choosing k distinct vectors from among the n unit coordinate vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$. This is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

And, indeed, it takes $\binom{4}{2} = 6$ parameters to determine an alternating 2-tensor on $\mathbb{R}^4$. One consequence of this is that, on $\mathbb{R}^n$, there are no alternating tensors of degree greater than n (except the zero tensor).

In any case, let us return to examining the constants $a_{i_1 \dots i_k}$. Recall that a *permutation* σ of a set k objects is just a rearrangement of them. If σ is any such permutation, we have

$$a_{i_1 \dots i_k} = \pm a_{\sigma(i_1 \dots i_k)}.$$

Any permutation can be performed by doing a sequence of pairwise exchanges; the sign on the left-hand side of the above equation depends on whether the number of such exchanges required is odd or even.

Definition 9. The *sign* of a permutation σ, denoted by $\operatorname{sgn}\sigma$, is either $+1$ or -1, whichever appears as the coefficient on the left-hand side of the above equation. A permutation with sign $+1$ is called *even*, and one with sign -1 is called *odd*.

An important point to note here is that the sign of a permutation is an intrinsic property of the permutation, having nothing to do with tensors or forms.

Definition 10. Let S be an alternating k-tensor on $\mathbb{R}^n$, and T and alternating l-tensor. Their *exterior* or *wedge product* $S \wedge T$ is a new alternating tensor of degree $k + l$, defined by

$$(S \wedge T)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}) = \frac{1}{k!l!} \sum_{\substack{\text{all permutations } \sigma \\ \text{of } k+l \text{ objects}}} (\operatorname{sgn}\sigma)(S \otimes T)(\sigma(\mathbf{v}_1, \dots, \mathbf{v}_{k+l})).$$