Tensors

Edoardo Milotti

October 12, 2023

1 Introduction

Tensors are introduced starting from generic vector spaces and are then heavily used in special vector spaces associated with *manifolds* (to be defined later). In this handout, we discuss the properties of tensors, with much material taken from [1] ch. 4 (grayed text), and [2] ch. 1 (blue text).

Note that in this handout we use the Einstein notation only sparsely, and still use summations to emphasize sums.

2 Linear and multilinear functions

We consider first linear and multilinear functions in generic vector spaces.

2.1 Effect of coordinate transformations on expressions for scalar invariants

Consider the scalar product

$$s = s(\mathbf{v}) = \boldsymbol{\lambda} \cdot \mathbf{v} = \sum_n \lambda_n v^n \tag{1}$$

which is a scalar function $s = s(\mathbf{v})$ of the vector $\mathbf{v}$ and whose representation does not depend on the reference frame. While $\boldsymbol{\lambda} \cdot \mathbf{v}$ is an abstract representation of the scalar product, the expression $\sum_n \lambda_n v^n$ requires explicit coordinates computed in a specific frame of reference. In a different frame of reference, the representation of the $\mathbf{v}$ vector is a linear combination of the coordinates in the first frame:

$$v'^i = A^i_j v^j \tag{2}$$

where the coefficients $\{A^i_j\}$ are a matrix representation of the linear transformation. The inverse transformation has a similar matrix representation $\{A^{-1}{}^j_i\}$, such that

$$\sum_k A^{-1}{}^k_i A^j_k = \delta^j_i \tag{3}$$

This means that we can write the scalar product in the form

$$s = \sum_n \lambda_n v^n = \sum_{i,j} \lambda_i \delta^i_j v^j = \sum_{i,j,k} \lambda_i {A^{-1}}^k_i A^k_j v^j = \sum_k \lambda'_k v'^k \tag{4}$$

and the scalar product remains constant if the vector of linear coefficients of the transformations obeys the transformation rule

$$\lambda'_j = {A^{-1}}^j_i \lambda^i \tag{5}$$

when the vector obeys the transformation law

$$v'^j = A^j_i v^i \tag{6}$$

2.2 Multilinear functions

We can extend the results of section 2.1 to multilinear functions that return a scalar as follows:

$$f(\mathbf{x}, \ldots, \mathbf{z}) = T_{i,\ldots,k} x^i \ldots z^k \tag{7}$$

(on the l.h.s, the coordinate-free representation of the function, on the r.h.s., the representation in a given coordinate frame). Then, as in section 2.1, we find that the function is unaffected by a coordinate change if

$$T'_{i,\ldots,k} = {A^{-1}}^j_i \ldots {A^{-1}}^\ell_k T_{j,\ldots,\ell} \tag{8}$$

A set of numbers $T_{i,\ldots,k}$ that transforms according to the rule (8) is called a *tensor*.

Tensors are important in many areas of physics, ranging from topics such as general relativity and electrodynamics to descriptions of the properties of bulk matter such as stress (the pattern of force applied to a sample) and strain (its response to the force), or the moment of inertia (the relation between a torsional force applied to an object and its resultant angular acceleration). Tensors constitute a generalization of other quantities: scalars and vectors. A scalar is quantity that remains invariant under rotations of the coordinate system and which can be specified by the value of a single real number. Vectors are identified as quantities that have a number of real components equal to the dimension of the coordinate system, with the components transforming like the coordinates of a fixed point when a coordinate system is rotated. Calling scalars tensors of rank 0 and vectors tensors of rank 1, we identify a tensor of rank n in a d-dimensional space as an object with the following properties:

- It has components labeled by n indices, with each index assigned values from 1 through d, and therefore having a total of d^n components;
- The components transform in a specified manner under coordinate transformations.

The behavior under coordinate transformation is of central importance for tensor analysis and conforms both with the way in which mathematicians define linear spaces and with the physicist's notion that physical observables must not depend on the choice of coordinate frames.

3 Tensors on differentiable manifolds

Although we have not yet introduced the concept of differentiable manifold, we shall now move to a different view of tensors, with local properties that depend on a generalization of the concept of parameterized continuous surface. We start with an example, provided by a spherical coordinate system used to map a sphere (see Figure 1).

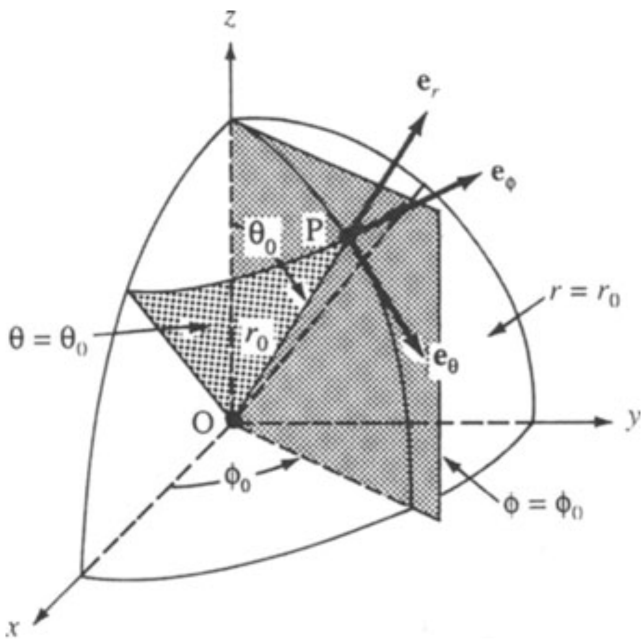

Figure 1: The coordinate surfaces and coordinate curves of spherical coordinates.

Suppose then that we have an alternate coordinate system (u, v, w) that is non-Cartesian, such as spherical coordinates (r, θ, ϕ), as in the example below. We can express the Cartesian coordinates x, y, z in terms of (u, v, w),

$$x = x(u, v, w); \quad y = y(u, v, w); \quad z = z(u, v, w), \tag{9}$$

and, in principle, invert these to get (u, v, w) in terms of x, y, z. Through any point P with coordinates (u, v, w) there pass three coordinate surfaces, given by $u = u_0$, $v = v_0$, and $w = w_0$, which meet in coordinate curves.

For spherical coordinates we have:

$$x = r \sin\theta \cos\phi, \quad y = r \sin\theta \sin\phi, \quad z = r \cos\theta, \tag{10}$$

where the conventional ranges for the coordinates are

$$r \geq 0, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi. \tag{11}$$

The coordinate surface $r = r_0$ is a sphere of radius r_0 (because $x^2 + y^2 + z^2 = r2$), the coordinate surface $\theta = \theta_0$ is an infinite cone with its vertex at the origin and its axis vertical, and the coordinate surface $\phi = \phi_0$ is a semi-infinite plane with the z axis as its edge. The surfaces $\theta = \theta_0$ and $\phi = \phi_0$ intersect to give a coordinate curve which is a ray (part of a line) that emanates from 0 and passes through P; the surfaces $\phi = \phi_0$ and $r = r_0$ intersect to give a coordinate curve which is a semicircle having its endpoints on the z axis and passing through P; and the surfaces $r = r_0$ and $\theta = \theta_0$ intersect to give a coordinate curve which is a horizontal circle passing through P with its center on the z axis.

The three equations (9) can be combined into a single vector equation that gives the position vector r of points in space as a function of the coordinates u, v, w that label the points:

$$\mathbf{r} = x(u, v, w)\hat{\mathbf{e}}_x + y(u, v, w)\hat{\mathbf{e}}_y + z(u, v, w)\hat{\mathbf{e}}_z, \tag{12}$$

where the $\hat{\mathbf{e}}$'s define an orthonormal cartesian reference frame. Setting w equal to the constant w_0, but leaving u, v to vary, gives

$$\mathbf{r} = x(u, v, w_0)\hat{\mathbf{e}}_x + y(u, v, w_0)\hat{\mathbf{e}}_y + z(u, v, w_0)\hat{\mathbf{e}}_z, \tag{13}$$

which is a parametric equation for the coordinate surface $w = w_0$ in which the coordinates u, v play the role of parameters. Parametric equations for the other two coordinate surfaces arise similarly. If we set $v = v_0$ and $w = w_0$, but let u vary, we get

$$\mathbf{r} = x(u, v_0, w_0)\hat{\mathbf{e}}_x + y(u, v_{,0}\, w_0)\hat{\mathbf{e}}_y + z(u, v_0, w_0)\hat{\mathbf{e}}_z, \tag{14}$$

which is a parametric equation for the coordinate curve given by the intersection of $v = v_0$ and $w = w_0$, in which the coordinate u acts as a parameter along the curve. Parametric equations for the other two coordinate curves arise similarly.

If we differentiate equation (12) with respect to the parameter u then we get a tangent vector to the coordinate curve. Since this differentiation is done holding v and w constant ($v = v_0$ and $w = w_0$), it amounts to differentiating equation (12) partially with respect to u. Similarly, by differentiating equation (12) partially with respect to v and w, we get tangent vectors to the other two coordinate curves. Thus, the three partial derivatives

$$\mathbf{e}_u = \frac{\partial \mathbf{r}}{\partial u}, \quad \mathbf{e}_v = \frac{\partial \mathbf{r}}{\partial v}, \quad \mathbf{e}_v = \frac{\partial \mathbf{r}}{\partial w} \tag{15}$$

when evaluated at (u_0, v_0, w_0), give tangent vectors to the three coordinate curves that pass through P.

The usual way forward in vector calculus is made on the assumption that the coordinate system is orthogonal (which means that the coordinate surfaces intersect orthogonally, so that the three vectors (15) are mutually orthogonal) and involves normalizing the vectors by dividing them by their lengths to get unit vectors. Thus

$$\hat{\mathbf{e}}_u = \frac{\mathbf{e}_u}{|\mathbf{e}_u|}, \quad \hat{\mathbf{e}}_v = \frac{\mathbf{e}_v}{|\mathbf{e}_v|}, \quad \hat{\mathbf{e}}_w = \frac{\mathbf{e}_w}{|\mathbf{e}_w|} \tag{16}$$

and these vector can be used as a local basis. However, our way forward does not require the coordinate system to be orthogonal, nor do we bother normalizing the tangent vectors to make them unit vectors. So at each point P we have the natural basis $(\mathbf{e}_u, \mathbf{e}_v, \mathbf{e}_w)$ produced by the partial derivative sand, in general, these vectors are neither unit vectors nor mutually orthogonal.

There is, in fact, another way in which the coordinate system (u, v, w) can be used to construct a basis at P. This uses the normals to the coordinate surfaces rather than the tangents to the coordinate curves. As remarked above, we can in principle invert equations (9) to obtain u, v, w in terms of x, y, z:

$$u = u(x, y, z), \quad v = v(x, y, z), \quad w = w(x, y, z). \tag{17}$$

This allows us to regard each coordinate as a scalar field and to calculate their gradients:

$$\begin{aligned} \nabla u &= \frac{\partial u}{\partial x}\hat{\mathbf{e}}_x + \frac{\partial u}{\partial y}\hat{\mathbf{e}}_y + \frac{\partial u}{\partial z}\hat{\mathbf{e}}_z \\ \nabla u &= \frac{\partial v}{\partial x}\hat{\mathbf{e}}_x + \frac{\partial v}{\partial y}\hat{\mathbf{e}}_y + \frac{\partial v}{\partial z}\hat{\mathbf{e}}_z \\ \nabla u &= \frac{\partial w}{\partial x}\hat{\mathbf{e}}_x + \frac{\partial w}{\partial y}\hat{\mathbf{e}}_y + \frac{\partial w}{\partial z}\hat{\mathbf{e}}_z \end{aligned} \tag{18}$$

At each point P, these gradient vectors are normal to the corresponding level surfaces through P, which are the coordinate surfaces $u = u_0, v = v_0, w = w_0$. We therefore obtain $(\nabla u, \nabla v, \nabla w)$ as an alternate basis at P. This basis is the dual of that obtained by using the tangent vectors to the coordinate curves and, to distinguish it from the previous one, we write its basis vectors with their suffixes as superscripts:

$$\hat{\mathbf{e}}^u = \nabla u, \quad \hat{\mathbf{e}}^v = \nabla v, \quad \hat{\mathbf{e}}^w = \nabla w. \tag{19}$$

Placing the suffixes in this position may seem odd at first (not least because of a possible confusion with powers), but it is part of a remarkably elegant and compact notation that will be developed more fully later. If the coordinate system is orthogonal, then the normals to the coordinate surfaces coincide with the tangents to the coordinate curves, making any distinction between $(\hat{\mathbf{e}}_u, \hat{\mathbf{e}}_v, \hat{\mathbf{e}}_w)$ and its dual $(\hat{\mathbf{e}}^u, \hat{\mathbf{e}}^v, \hat{\mathbf{e}}^w)$ just a matter of lengths, rather than the lengths and directions of the basis vectors. If the basis vectors are normalized, then the distinction disappears altogether. Consequently, to illustrate better the two bases that arise naturally from the coordinate system, we should use one that is not orthogonal, rather than continue using spherical coordinates.

4 Covariant and contravariant tensors

When we consider the rotational transformation of a vector $\mathbf{v} = v_1\hat{\mathbf{e}}_1 + v_2\hat{\mathbf{e}}_2 + v_3\hat{\mathbf{e}}_3$ from the (for the moment ...) Cartesian system defined by $\hat{\mathbf{e}}_i$, $(i = 1, 2, 3)$ into a rotated coordinate system defined by $\hat{\mathbf{e}}'_i$, with the same vector $\mathbf{v}$ then represented as $\mathbf{v} = v'_1\hat{\mathbf{e}}'_1 + v'_2\hat{\mathbf{e}}'_2 + v'_3\hat{\mathbf{e}}'_3$. The two representations are related by

$$v'_i = \sum_j (\hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j) v_j \tag{20}$$

We temporarily limit this discussion to Cartesian systems, and we recall that in the case of a sphere we could define a translated reference frame such that

$$\hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j = \frac{\partial x'_i}{\partial x_j} \tag{21}$$

The r.h.s. corresponds to the application of the chain rule to convert the set v_j into the set v'_i, and is valid for v_j and v'_i of arbitrary magnitude because both vectors depend linearly on their components.

We note that the gradient of a scalar ϕ has in the unrotated Cartesian coordinates the components

$$(\nabla\phi)_i = \frac{\partial \phi}{\partial x_i}\hat{\mathbf{e}}_i$$

meaning that in a rotated system we would have

$$(\nabla\phi)'_i = \frac{\partial \phi}{\partial x'_i} = \sum_j \frac{\partial x_j}{\partial x'_i}\frac{\partial \phi}{\partial x_j} \tag{22}$$

showing that the gradient has a transformation law that differs from that of Eq. (20) in that $\partial x'_i/\partial x_j$ has been replaced by $\partial x_j/\partial x'_i$. Here we note, as the alert reader may note from the repeated insertion of the word "Cartesian", that the partial derivatives in (20) and (22) are only guaranteed to be equal in Cartesian coordinate systems, and since there is sometimes a need to use non-Cartesian systems it becomes necessary to distinguish these two different transformation rules. Quantities transforming according to Eq. (20) are called contravariant vectors, while those transforming according to Eq. (22) are termed covariant. **When non-Cartesian systems may be in play, it is therefore customary to distinguish these transformation properties by writing the index of a contravariant vector as a superscript and that of a covariant vector as a subscript**. This means, among other things, that the components of the position vector $\mathbf{r}$, which are contravariant, must now be written (x^1, x^2, x^3). Thus, summarizing,

$$v'^i = \sum_j \frac{\partial x'^i}{\partial x^j} v^j; \quad \text{for a contravariant vector } \{v^i\} \tag{23a}$$

$$v'_i = \sum_j \frac{\partial x^j}{\partial x'^i} v_j; \quad \text{for a covariant vector } \{v_i\} \tag{23b}$$

It is useful to note that the occurrence of subscripts and superscripts is systematic; the free (i.e., unsummed) index i occurs as a superscript on both sides of Eq. (23a), while it appears as a subscript on both sides of Eq. (23b), if we interpret an upper index in the denominator as equivalent to a lower index. The summed index occurs once as upper and once as lower (again treating an upper index in the denominator as a lower index). A frequently used shorthand (the Einstein convention) is to omit the summation sign in formulas like Eqs. (23a) and (23b) and to understand that when the same symbol occurs both as an upper and a lower index in the same expression, it is to be summed.

5 Tensors of rank 2

Now we proceed to define contravariant, mixed, and covariant tensors of rank 2 by the following equations for their components under coordinate transformations:

$$
\begin{aligned}
A'^{ij} &= \frac{\partial x'^i}{\partial x^k}\,\frac{\partial x'^j}{\partial x^\ell} A^{k\ell} \\
B'^{i}_{j} &= \frac{\partial x'^i}{\partial x^k}\,\frac{\partial x^\ell}{\partial x'^j} B^{k}_{\ell} \\
C'_{ij} &= \frac{\partial x^k}{\partial x'^i}\,\frac{\partial x^\ell}{\partial x'^j} C_{k\ell}
\end{aligned} \tag{24}
$$

Clearly, the rank goes as the number of partial derivatives (or direction cosines) in the definition: 0 for a scalar, 1 for a vector, 2 for a second-rank tensor, and so on. Each index (subscript or superscript) ranges over the number of dimensions of the space. The number of indices (equal to the rank of tensor) is not limited by the dimensionality of the space. We see that $A^{k\ell}$ is contravariant with respect to both indices, $C_{k\ell}$ is covariant with respect to both indices, and B^k_ℓ transforms contravariantly with respect to the index k but covariantly with respect to the index ℓ. Once again, if we are using Cartesian coordinates, all three forms of the tensors of second rank, contravariant, mixed, and covariant are the same. **As with the components of a vector, the transformation laws for the components of a tensor, Eq.** (24)**, cause its physically relevant properties to be independent of the choice of reference frame. This is what makes tensor analysis important in physics. The independence relative to reference frame (invariance) is ideal for expressing and investigating universal physical laws**.

In summary, tensors are systems of components organized by one or more indices that transform according to specific rules under a set of transformations. The number of indices is called the rank of the tensor.

6 Addition and subtraction of tensors

The addition and subtraction of tensors is defined in terms of the individual elements, just as for vectors. For instance, for rank 2 contravariant tensors

$$A = B + C \tag{25}$$

means

$$A^{ij} = B^{ij} + C^{ij} \tag{26}$$

in terms of components. In general, of course, A and B must be tensors of the same rank (of both contra- and co-variance) and in the same space.

7 Symmetry

The order in which the indices appear in our description of a tensor is important. In general, A^{ij} is independent of A^{ji} , but there are some cases of special interest. If, for all i and j,

$$A^{ij} = A^{ji} \quad \Rightarrow \quad \text{A is symmetric.} \tag{27}$$

If, on the other hand,

$$A^{ij} = -A^{ji} \quad \Rightarrow \quad \text{A is antisymmetric.} \tag{28}$$

Clearly, every (second-rank) tensor can be resolved into symmetric and antisymmetric parts by the identity

$$A^{ij} = \frac{1}{2}\left(A^{ij} + A^{ji}\right) + \frac{1}{2}\left(A^{ij} - A^{ji}\right) \tag{29}$$

the first term on the right being a symmetric tensor, the second, an antisymmetric tensor.

8 Kroneker's delta and isotropic tensors

To illustrate some of the techniques of tensor analysis, let us show that the now-familiar Kronecker delta, $\delta_{k\ell}$, is really a mixed tensor of rank 2, δ^k_ℓ. The question is: Does δ^k_ℓ transform according to Eq. (24)? This is our criterion for calling it a tensor. If δ^k_ℓ is the mixed tensor corresponding to this notation, it must satisfy

$$\delta'^i_j = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^\ell}{\partial x'^j} \delta^k_\ell, \tag{30}$$

and indeed

$$\frac{\partial x'^i}{\partial x^k} \frac{\partial x^\ell}{\partial x'^j} \delta^k_\ell = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x'^j} = \frac{\partial x'^i}{\partial x'^j} = \delta'^i_j, \tag{31}$$

showing that the δ^k_ℓ are indeed the components of a mixed second-rank tensor. Note that this result is independent of the number of dimensions of our space. **The Kronecker delta has one further interesting property. It has the same components in all of our rotated coordinate systems and is therefore called isotropic.**

9 Contraction

When dealing with vectors, we form a scalar product by summing products of corresponding components:

$$\mathbf{a} \cdot \mathbf{b} = \sum_i a_i b_i. \tag{32}$$

The generalization of this expression in tensor analysis is a process known as contraction. Two indices, one covariant and the other contravariant, are set equal to each other, and then (as implied by the summation convention) we sum over this repeated index. For example, let us contract the second-rank mixed tensor Bi_j by setting $i = j$, then summing over i. To see what happens, let's look at the transformation formula that converts B into B':

$$B'^i_i = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^\ell}{\partial x'^i} B^k_\ell = \frac{\partial x^\ell}{\partial x^k} B^k_\ell = \delta^\ell_k B^k_\ell = B^k_k. \tag{33}$$

We see that the contracted B is invariant under transformation and is therefore a scalar; in matrix analysis this scalar is the *trace* of the matrix. In general, the operation of contraction reduces the rank of a tensor by 2.

10 Direct product

The components of two tensors (of any ranks and covariant/contravariant characters) can be multiplied, component by component, to make an object with all the indices of both factors. The new quantity, termed the direct product of the two tensors, can immediately be shown to be a tensor whose rank is the sum of the ranks of the factors, and with covariant/contravariant character that is the sum of those of the factors. For example:

$$C^{ij}_{klm} = A^i_k B^j_{lm}, \quad F^{ij}_{kl} = A^i B^j_{kl}. \tag{34}$$

Note that the way we take the index groups from different tensors in the direct product is not important, but the covariance/contravariance of the factors must be maintained in the direct product. Note also that the direct product concept gives a meaning to quantities such as $\nabla\mathbf{E}$, which is not defined within the framework of vector analysis. However, this and other tensor-like quantities involving differential operators must be used with caution, because their transformation rules are simple only in Cartesian coordinate systems. In non-Cartesian systems, operators $\partial/\partial x_i$ act also on the partial derivatives in the transformation expressions and alter the tensor transformation rules.

11 Inverse transformation

If we have a contravariant vector a^i , which must have the transformation rule

$$a'^j = \frac{\partial x'^j}{\partial x^k} a^k \tag{35}$$

the inverse transformation (which can be obtained simply by interchanging the roles of the primed and unprimed quantities) is

$$a^j = \frac{\partial x^j}{\partial x'^k} a'^k \tag{36}$$

as may also be verified as follows

$$\frac{\partial x^i}{\partial x'^j} a'^j = \frac{\partial x^i}{\partial x'^j} \frac{\partial x'^j}{\partial x^k} a^k = \frac{\partial x^i}{\partial x^k} a^k = \delta^i_k a^k = a^i. \tag{37}$$

We see that a^i is recovered. Incidentally, note that the inverse does not propagate to individual tensor elements

$$\left(\frac{\partial x'^j}{\partial x^i}\right)^{-1} \neq \frac{\partial x^i}{\partial x'^j} \tag{38}$$

these derivatives have different other variables held fixed. The cancellation in Eq. (37) only occurs because the product of derivatives is summed. In Cartesian systems, we do have

$$\frac{\partial x^i}{\partial x'^j} = \frac{\partial x'^j}{\partial x^i} \tag{39}$$

(the transpose of a rotation matrix is equal to the inverse) both equal to the direction cosine connecting the x^i and x'^j axes, but this equality does not extend to non-Cartesian systems.

Note that the matrix elements of these transformation matrices define the *Jacobian matrices*, and what we have just found is that Jacobian matrices for inverse transformations are the inverse of each other.

12 Quotient rule

If, for example, A_{ij} and B_{kl} are tensors, we have already observed that their direct product, $A_{ij}B_{kl}$, is also a tensor. Here we are concerned with the inverse problem, illustrated by equations such as

$$\begin{aligned} K_i A^i &= B, \\ K_i^j A_j &= B_i \\ K_i^j A_{jk} &= B_{ik} \\ K_{ijkl} A^{ij} &= B_{kl} \\ K^{ij} A^k &= B^{ijk} \end{aligned} \tag{40}$$

In each of these expressions A and B are known tensors of ranks indicated by the number of indices. In each case K is an unknown quantity. We wish to establish the transformation properties of K. The quotient rule asserts: *If the equation of interest holds in all transformed coordinate systems, then K is a tensor of the indicated rank and* *covariant/contravariant character*. Part of the importance of this rule in physical theory is that it can establish the tensor nature of quantities. For example, the equation giving the dipole moment **p** induced in an anisotropic medium by an electric field **E** is

$$p_i = P_{ij}E^j. \tag{41}$$

Since we know that **p** and **E** are vectors, the general validity of this equation tells us that the polarization matrix P is a tensor of rank 2.

Let's prove the quotient rule for a typical case, which we choose to be the second of Eqs. (40). If we apply a transformation to that equation, we have

$$K_i^j A_j = B_i \quad \rightarrow \quad K'^j_i A'_j = B'_i \tag{42}$$

and when we write A' and B' using the transformation formulas, we find

$$K'^j_i \frac{\partial x^k}{\partial x'^j} A_k = \frac{\partial x^l}{\partial x'^i} B_l, \tag{43}$$

i.e.,

$$\frac{\partial x'^i}{\partial x^l} \frac{\partial x^k}{\partial x'^j} K'^j_i A_k = B_l, \tag{44}$$

Comparing this with the original expression, and taking into account the arbitrariness of A, we find

$$K_l^k = \frac{\partial x'^i}{\partial x^l} \frac{\partial x^k}{\partial x'^j} K'^j_i \tag{45}$$

i.e.,

$$K'^j_i = \frac{\partial x^l}{\partial x'^i} \frac{\partial x'^j}{\partial x^k} K_l^k \tag{46}$$

which is the expected transformation for a mixed tensor of rank 2. Other cases may be treated similarly. One minor pitfall should be noted: The quotient rule does not necessarily apply if B is zero. The transformation properties of zero are indeterminate.

13 Pseudotensors and dual tensors

The topics of this section will be treated for tensors restricted for practical reasons to Cartesian coordinate systems. This restriction is not conceptually necessary but simplifies the discussion and makes the essential points easy to identify.

Here, we consider the effect of reflections or inversions of the coordinate system (sometimes also called **improper rotations**). Restricting our attention to orthogonal systems of Cartesian coordinates, we see that the effect of a coordinate rotation on a fixed vector can be described by a transformation of its components according to the formula

$$\mathbf{v}' = S\mathbf{v} \tag{47}$$

where S is an orthogonal matrix with determinant $+1$. If the coordinate transformation includes a reflection (or inversion), the transformation matrix is still orthogonal, but has determinant -1. While this transformation rule is obeyed by vectors describing quantities such as position in space or velocity, it produces the wrong sign when vectors describing angular velocity, torque, and angular momentum are subject to improper rotations. These quantities, called axial vectors, or nowadays pseudovectors, obey the transformation rule

$$\mathbf{v}' = \det(S)\ S\mathbf{v} \quad \text{(pseudovector)} \tag{48}$$

The extension of this concept to tensors is straightforward. We consider the possibility of having, at arbitrary rank, objects whose transformation is like that of tensors, but requires an additional sign factor to adjust for the effect associated with improper rotations. These objects are called pseudotensors, and constitute a generalization of the objects already identified as pseudoscalars and pseudovectors.

13.1 The Levi-Civita symbol

The three-index version of the Levi-Civita symbol has the values

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1 \tag{49}$$

$$\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1 \tag{50}$$

$$\text{all other } \epsilon_{ijk} = 0 \tag{51}$$

Suppose now that we have a rank-3 pseudotensor η_{ijk} , which in one particular Cartesian coordinate system is equal to ϵ_{ijk}. Then, letting A stand for the matrix of coefficients $\{a_{ij}\}$ in an orthogonal transformation of $\mathbb{R}^3$, we have in the transformed coordinate system

$$\eta'_{ijk} = \det(A) \sum_{lmn} a_{il} a_{jm} a_{kn} \epsilon_{lmn} \tag{52}$$

by definition of pseudotensor. All terms of the pqr sum will vanish except those where pqr is a permutation of 123, and when pqr is such a permutation the sum will correspond to the determinant of A except that its rows will have been permuted from 123 to ijk . This means that the pqr sum will have the value $\det(A)\epsilon_{ijk}$, and

$$\eta'_{ijk} = [\det(A)]^2 \epsilon_{ijk} = \epsilon_{ijk}, \tag{53}$$

where the final result depends on the fact that $|\det(A)| = 1$.

Equation (53) not only shows that ϵ is a rank-3 pseudotensor, but also that it is also isotropic. In other words, it has the same components in all rotated Cartesian coordinate systems, and -1 times those component values in all Cartesian systems that are reached by improper rotations.

13.2 Dual tensors

With any **antisymmetric** second-rank tensor C (in 3-D space) we may associate a pseudovector **C** with components defined by

$$C_i = \frac{1}{2}\epsilon_{ijk}C^{jk}. \tag{54}$$

In matrix form the antisymmetric C may be written

$$C = \begin{pmatrix} 0 & C^{12} & -C^{31} \\ -C^{12} & 0 & C^{23} \\ C^{31} & -C^{23} & 0 \end{pmatrix} \tag{55}$$

We know that C_i must transform as a vector under rotations because it was obtained from the double contraction $\epsilon_{ijk}C^{jk}$, but that it is really a pseudovector because of the pseudo nature of ϵ_{ijk} . Specifically, the components of C are given by

$$(C^1, C^2, C^3) = (C^{23}, C^{31}, C^{12}). \tag{56}$$

Note the cyclic order of the indices that comes from the cyclic order of the components of ϵ_{ijk}. We identify the pseudovector of Eq. (56) and the antisymmetric tensor of Eq. (55) as dual tensors; they are simply different representations of the same information. Which of the dual pair we choose to use is a matter of convenience.

Here is another example of duality. If we take three vectors A, B, and C, we may define the direct product

$$V^{ijk} = A^i B^j C^k \tag{57}$$

which is evidently a rank-3 tensor. The dual quantity

$$V = \epsilon_{ijk}V^{ijk} \tag{58}$$

is clearly a pseudoscalar. By expansion it is seen that

$$\begin{vmatrix} A^1 & B^1 & C^1 \\ A^2 & B^2 & C^2 \\ A^3 & B^3 & C^3 \end{vmatrix},$$

the familiar scalar triple product.

References

[1] GB Arfken, HJ Weber, and FE Harris. *Mathematical Methods for Physicists: A Comprehensive Guide. 7th Edition.* Academic Press, 2012.

[2] James Foster, J David Nightingale, and J Foster. *A short course in General Relativity.* Springer, 1995.