Physics 114A
Tensors

Peter Young

Contents

I. INTRODUCTION

In elementary classes you met the concept of a *scalar*, which is described as something with a magnitude but no direction, and a *vector* which is intuitively described has having both direction and magnitude. In this part of the course we will:

1. Give a precise meaning to the intuitive notion that a vector "has both direction and magnitude."

2. Realize that there are more general quantities, also important in physics, called tensors, of which scalars and vectors form two classes.

Most of the handout will involve somewhat formal manipulations. In Sec. VI we will discuss the main usefulness of tensor analysis in physics. Tensors are particularly important in special and general relativity.

For most of this handout will will discuss *cartesian tensors* which in which we consider how things transform under ordinary rotations. However, in Sec. VII we will discuss tensors which involve the Lorentz transformation in special relativity. In this handout we will not discuss more general tensors which are needed for general relativity.

After a rotation, the coordinates,[1] x_1, x_2, x_3, of a point $\overrightarrow{x}$ become x'_1, x'_2, x'_3, where

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = U \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{1}$$

and U is a 3×3 rotation matrix. In terms of components this can be written

$$x'_i = \sum_{j=1}^{3} U_{ij} x_j. \tag{2}$$

Note that the repeated index j is summed over. This happens so often that we will, from now on, follow the Einstein convention of *not* writing explicitly the summation over a repeated index, so Eq. (2) will be expressed as

$$x'_i = U_{ij} x_j. \tag{3}$$

Let us emphasize that Eqs. (2) and (3) mean exactly the same thing, but Eq. (3) is more compact and "elegant".

In order that the matrix U represents a rotation without any stretching or "shearing" it must be orthogonal. This is proved in Appendix A. A 3×3 real orthogonal matrix has three independent parameters.[2] These are often taken to be the three Euler angles, defined, for example in Arfken and Weber, p. 188–189.

Euler angles are a bit complicated so, to keep things simple, we will here restrict ourselves to *rotations in a plane*. Hence vectors have just 2 components and U is a 2×2 matrix. Clearly there is just one angle involved,[3] the size of the rotation about an axis perpendicular to the plane, see Fig. 1.

As shown in class and in Arfken and Weber p. 185, the relation between the primed and the unprimed components is

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{4}$$

so

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{5}$$

In terms of components we have

$$U_{11} = U_{22} = \cos\theta, \qquad U_{12} = -U_{21} = \sin\theta. \tag{6}$$

---

[1] We prefer the symbols $x_i, i = 1, 2, 3$ rather than x, y and z, because (i) we can generalize to an arbitrary number of components, and (ii) we can use the convenient summation symbol to sum over components.

[2] To see this note that there are 9 elements altogether, but there are 6 constraints: 3 coming from each column being normalized, and $3(3-1)/2 = 3$ more coming from distinct columns having to be orthogonal. Hence the number of independent parameters is $9 - 6 = 3$.

[3] If you want to see this mathematically using similar reason to that in footnote 2 for 3 components, note that a 2×2 real orthogonal matrix has four elements but there are three constraints, since there are two column vectors which must be normalized and $2(2-1)/2 = 1$ pairs of distinct column vectors which must be orthogonal. Now $4 - 3 = 1$, so there is just one parameter as expected.

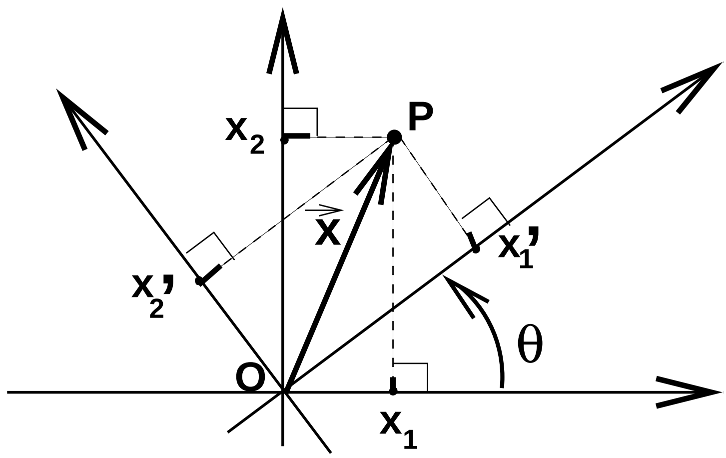

FIG. 1: A rotation in a plane through an angle θ. The components of a point $\vec{x}$ can either be expressed relative to the original or rotated (primed) axes. The connection between the two is given by Eq. (4).

II. WHAT IS A VECTOR?

We are now in a position to give a *precise* definition of a (cartesian) vector:

> A quantity $\vec{A}$ is a vector if its components, A_i, transform into each other under rotations in the same way as the components of position, x_i.

In other words, if

$$x_i' = U_{ij} x_j, \tag{7}$$

where U describes a rotation, then the components of $\vec{A}$ in the rotated coordinates must be given by

$$A_i' = U_{ij} A_j. \tag{8}$$

This is the precise meaning of the more vague concept of "a quantity with direction and magnitude". For two components, the transformation law is

$$\begin{aligned} A_1' &= \cos\theta\ A_1 + \sin\theta\ A_2 \\ A_2' &= -\sin\theta\ A_1 + \cos\theta\ A_2. \end{aligned} \tag{9}$$

Similarly, a scalar is a quantity which is *invariant* under a rotation of the coordinates.

It is easy to see that familiar quantities such as velocity and momentum are vectors according to this definition, as we expect. Since time is a scalar[4] (it does not change in a rotated coordinate system), then clearly the components of velocity v_i, defined by

$$v_i = \frac{dx_i}{dt}, \tag{10}$$

[4] Note that time is not a scalar in special relativity where we consider Lorentz transformations as well as rotations. We will discuss this later. However, for the moment, we are talking about cartesian tensors, where we are only interested in the transformation properties under rotations.

transform under rotations in the same way as the components of position. Velocity is therefore a vector. Also, the mass, m, of an object is a scalar[5] so the momentum components, given by

$$p_i = m v_i, \tag{11}$$

transform in the same way as those of the velocity, (and hence in the same way as those of position), so momentum is also a vector.

Now we know that

$$\begin{pmatrix} x \\ y \end{pmatrix} \tag{12}$$

is a vector but what about

$$\begin{pmatrix} -y \\ x \end{pmatrix} \quad , \qquad \text{i.e.} \quad \begin{array}{l} A_1 = -y \\ A_2 = \ \ x \end{array} \quad ? \tag{13}$$

Under rotations we have

$$\begin{aligned} A_1 \to A_1' &= -y' \\ &= -(\cos\theta\ y - \sin\theta\ x) \\ &= \cos\theta\ A_1 + \sin\theta\ A_2. \end{aligned} \tag{14}$$

Similarly

$$\begin{aligned} A_2 \to A_2' &= x' \\ &= (\cos\theta\ x + \sin\theta\ y) \\ &= \cos\theta\ A_2 - \sin\theta\ A_1. \end{aligned} \tag{15}$$

Eqs. (14) and (15) are just the desired transformation properties of a vector, see Eq. (9), hence Eq. (13) is a vector. Physically, we started with a vector (x, y) and rotated it by 90° (which still leaves it a vector) to get Eq. (13).

However, if we try

$$\begin{pmatrix} y \\ x \end{pmatrix} \quad , \qquad \text{i.e.} \quad \begin{array}{l} A_1 = y \\ A_2 = x \end{array}, \tag{16}$$

then

$$\begin{aligned} A_1 \to A_1' &= y' \\ &= (\cos\theta\ y - \sin\theta\ x) \\ &= \cos\theta\ A_1 - \sin\theta\ A_2, \end{aligned} \tag{17}$$

$$\begin{aligned} A_2 \to A_2' &= x' \\ &= (\cos\theta\ x + \sin\theta y) \\ &= \cos\theta\ A_2 + \sin\theta\ A_1. \end{aligned} \tag{18}$$

Eqs. (17) and (18) do *not* correspond to the transformation of vector components in Eq. (9). Hence Eq. (16) is not a vector.

[5] This also needs further discussion in special relativity.

Let us emphasize that a set of quantities with a subscript, *e.g.* A_i, is *not* necessarily a vector. One has to show that the components transform into each other under rotations in the desired manner.

Another quantity that we have just assumed in the past to be a vector is the gradient of a scalar function. In Appendix B we show that it really is a vector because U is orthogonal, *i.e.* $\left(U^{-1}\right)^T = U$. However, in Sec. VII we will discuss other types of transformations, in particular Lorentz transformations, for which $\left(U^{-1}\right)^T \neq U$. We shall then need to define *two* types of vectors, one transforming like the x_i and the other transforming like the derivatives $\partial\phi/\partial x_i$, where ϕ is a scalar function. The transformation properties of derivatives under more general transformations are also discussed in Appendix B.

III. WHAT IS A TENSOR?

We have seen that a *scalar* is a quantity with no indices that does not change under a rotation, and that a *vector* is a set of quantities, labeled by a single index, which transform into each other in a specified way under a rotation. There are also quantities of importance in physics which have more than one index and transform into each other in a more complicated way, to be defined below. These are called *tensors*. If there are n indices we say that the tensor is of rank n. In fact a vector is a special case, namely a tensor of rank one, and a scalar is a tensor of rank 0. Firstly I will give an example of a second rank tensor, and then state the transformation properties of tensors.

Consider an object of mass m at position $\vec{x}$ moving with velocity $\vec{v}$. The angular velocity, $\vec{\omega}$, is related to these in the usual way:

$$\vec{v} = \vec{\omega} \times \vec{x}. \tag{19}$$

The angular momentum[6] is given by

$$\begin{aligned} \vec{L} &= m\vec{x} \times \vec{v} \\ &= m\vec{x} \times (\vec{\omega} \times \vec{x}) \\ &= m\left(x^2\vec{\omega} - (\vec{x}\cdot\vec{\omega})\,\vec{x}\right), \end{aligned} \tag{20}$$

where $x^2 \equiv x_k x_k$, and the last line uses the expression for a triple vector product

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \left(\vec{A}\cdot\vec{C}\right)\vec{B} - \left(\vec{A}\cdot\vec{B}\right)\vec{C}. \tag{21}$$

Eq. (20) can be expressed in components as

$$L_i = m\left[x^2\omega_i - x_j\omega_j\, x_i\right] \tag{22}$$

$$\quad = I_{ij}\omega_j, \tag{23}$$

[6] Remember the rate of change of angular momentum is equal to the torque. Thus, angular momentum plays, for rotational motion, the same role that ordinary momentum plays for translational motion.

where repeated indices are summed over, and I_{ij}, the moment of inertia[7], is given by

$$I_{ij} = m\left[x^2\delta_{ij} - x_i x_j\right]. \tag{24}$$

We shall see that the moment of inertia is an example of a second rank tensor.

Typically one is interested in the motion of inertia of a rigid body rather than an single point mass. To get the moment of inertia of a rigid body one takes Eq. (24), replaces m by the density ρ, and integrates over the body.

Note that Eq. (23) is the analogue, for rotational motion, of the familiar equation for translational motion, $\overrightarrow{p} = m\overrightarrow{v}$, where $\overrightarrow{p}$ is the momentum, which can be written in component notation as

$$p_i = m v_i. \tag{25}$$

An important difference between translational and rotational motion is that, since m is a scalar, $\overrightarrow{p}$ and $\overrightarrow{v}$ are always in the same direction, whereas, since I_{ij} is a second rank tensor, $\overrightarrow{L}$ and $\overrightarrow{\omega}$ are *not* necessarily in the same direction. This is one of the main reasons why rotational motion is more complicated and harder to understand than translational motion.

We define a second rank tensor, by analogy with a vector as follows:

> A second rank tensor is a set of quantities with two indices, T_{ij}, which transform into each other under a rotation as
>
> $$T'_{ij} = U_{ik}U_{jl}T_{kl}, \tag{26}$$
>
> *i.e.* **the components transform in the same way as the products** $x_i x_j$.

Note the pattern of indices in this last equation (which will persist for tensors of higher rank so you should remember it):

- There is an element of the matrix U for each index of the tensor T.
- The first index on each of the U-s is the same as one of the indices of the rotated tensor component on the left.
- The second index on each of the U-s is the same as an index of the unrotated tensor component on the right.

We can represent a second rank tensor as a matrix:

$$T = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}. \tag{27}$$

Note also that we can write Eq. (26) as

$$A'_{ij} = U_{ik}T_{kl}U^T_{lj}, \tag{28}$$

[7] It should be clear from the context whether I_{ij} refers to the moment of inertia tensor and when to the identity matrix

or, in matrix notation,

$$T' = UTU^T, \tag{29}$$

which, with a special choice of U, is the transformation which diagonalizes a real symmetric matrix (see the handout on diagonalization). Hence we see that the moment of inertia tensor will have a diagonal form (*i.e.* will only be non-zero if $i = j$) in a particular set of axes, called the *principal axes*.

We should emphasize that writing transformations in tensor notation is more general than matrix notation, Eq. (29), because it can be used for tensors of higher rank. The definition of tensors of higher rank follows in an obvious manner, *e.g.* a third rank tensor transforms in the same way as $x_i x_j x_k$, *i.e.*

$$T'_{ijk} = U_{il} U_{jm} U_{kn} T_{lmn}. \tag{30}$$

An n-th rank tensor transforms in the same way as $x_{i_1} x_{i_2} \cdots x_{i_n}$, *i.e.*

$$T'_{i_1 i_2 \cdots i_n} = U_{i_1 j_1} U_{i_2 j_2} \cdots U_{i_n j_n} T_{j_1 j_2 \cdots j_n}. \tag{31}$$

For $n > 2$ tensors cannot be represented as matrices.

For rotations in a plane, a second rank tensor has 4 components, T_{xx}, T_{xy}, T_{yx} and T_{yy}. From Eq. (26), T_{11} becomes, in the rotated coordinates,

$$T'_{11} = U_{11} U_{11} T_{11} + U_{11} U_{12} T_{12} + U_{12} U_{11} T_{21} + U_{12} U_{12} T_{22}, \tag{32}$$

(notice the pattern of the indices). From Eq. (5), this can be written explicitly as

$$T'_{11} = \cos^2\theta \; T_{11} + \sin\theta\cos\theta \; T_{12} + \sin\theta\cos\theta \; T_{21} + \sin^2\theta \; T_{22}. \tag{33}$$

Similarly one finds

$$\begin{aligned}
T'_{12} &= -\cos\theta\sin\theta \; T_{11} + \cos^2\theta \; T_{12} - \sin^2\theta \; T_{21} + \cos\theta\sin\theta \; T_{22} && (34)\\
T'_{21} &= -\cos\theta\sin\theta \; T_{11} - \sin^2\theta \; T_{12} + \cos^2\theta \; T_{21} + \cos\theta\sin\theta \; T_{22} && (35)\\
T'_{22} &= \sin^2\theta \; T_{11} - \cos\theta\sin\theta \; T_{12} - \cos\theta\sin\theta \; T_{21} + \cos^2\theta \; T_{22} && (36)\\
& && (37)
\end{aligned}$$

We can now verify that the components of the moment of inertia in Eq. (24) do form a tensor. From Eq. (24), the moment of inertia can be expressed in matrix notation as (in units where $m = 1$)

$$I = \begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix}, \qquad \text{i.e.} \quad \begin{matrix} I_{11} = y^2, & I_{12} = -xy \\ I_{21} = -xy, & I_{22} = x^2 \end{matrix}. \tag{38}$$

Hence, in the rotated coordinates,

$$\begin{aligned}
I'_{11} &= y'^2 \\
&= (-\sin\theta \; x + \cos\theta \; y)^2 \\
&= \sin^2\theta \; x^2 - \sin\theta\cos\theta \; xy - \sin\theta\cos\theta \; xy + \cos^2\theta \; y^2 \\
&= \cos^2\theta \; I_{11} + \sin\theta\cos\theta \; I_{12} + \sin\theta\cos\theta \; I_{21} + \sin^2\theta \; I_{22}, && (39)
\end{aligned}$$

which is indeed the correct transformation property given in Eq. (33). Repeating the calculation for the other three components gives the results in Eqs. (34)–(36). This is a bit laborious. Soon we will see a quicker way of showing that the moment of inertia is a second rank tensor.

If one repeats the above analysis for

$$\begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}, \tag{40}$$

where have just changed the sign of the off-diagonal elements, then one finds that the transformation properties are *not* given correctly, so this is not a second rank tensor.

We have already met a quantity with two indices, the Kronecker delta function, δ_{ij}, which is 1 if $i = j$ and zero otherwise *independent of any rotation*. Is this a tensor? As we have just seen in the last example, all quantities with two indices are not necessarily tensors, so we need to *show*, through its transformation properties, that δ_{ij} is a second rank tensor. Under a rotation, the Kronecker delta function becomes

$$\begin{aligned} \delta'_{ij} &= U_{ik}U_{jl}\delta_{kl} \\ &= U_{ik}U_{jk} = U_{ik}U^T_{kj} = \left(UU^T\right)_{ij} \\ &= \delta_{ij}, \end{aligned} \tag{41}$$

where the last line follows since U is orthogonal and hence $UU^T = UU^{-1} = I$, the identity matrix. This is just what we wanted: the delta function has the same properties independent of any rotation. Hence δ_{ij} is an *isotropic*[8] second rank tensor, *i.e.* transforming it according to the rules for a second rank tensor it is the *same in all rotated frames of reference*. Note that it is *not* a scalar because a scalar is a single number whereas the delta function has several elements.

Clearly the sum or difference of two tensors of the same rank is also a tensor, *e.g.* if A and B are second rank tensors then

$$C_{ij} = A_{ij} + B_{ij} \tag{42}$$

is a second rank tensor. Similarly if one multiplies all elements of a tensor by a scalar it is still a tensor, *e.g.*

$$C_{ij} = \lambda A_{ij} \tag{43}$$

is a second rank tensor. Also, multiplying a tensor of rank n with one of rank m gives a tensor of rank $n + m$, *e.g.* if A and B are third rank tensors then

$$C_{ijkmnp} = A_{ijk}B_{mnp} \tag{44}$$

is a 6-th rank tensor. To see this note that both sides of Eq. (44) transform like $x_i x_j x_k x_m x_n x_p$.

Frequently one can conveniently show that a quantity is a tensor by writing it in terms of quantities that we know are tensors. As an example let us consider again the moment of inertia tensor in Eq. (24). The second term, proportional to $x_i x_j$, is, *by definition* a second rank tensor,[9] and the first term is the product of a scalar, mx^2 (where x^2 is the square of the length of a vector which is invariant under rotation), and δ_{ij}, which we have just shown is a second rank tensor. Hence the moment of inertia *must* also be a second rank tensor. This is a much simpler derivation than the one above, where we verified explicitly that the transformation properties are correct.

[8] Also called an invariant tensor.

[9] See the discussion below Eq. (26).

A *symmetric* second rank tensor is defined to be one that satisfies

$$T_{ji} = T_{ij}, \tag{45}$$

and an *antisymmetric* second rank tensor is defined by

$$T_{ji} = -T_{ij}. \tag{46}$$

It is easy to see from the transformation properties that a symmetric tensor stays symmetric after rotation, and similarly for an antisymmetric tensor.

An example of a symmetric tensor is the moment of inertia tensor, see Eqs. (24) and Eq. (38), discussed earlier in this section. As an example of an antisymmetric second rank tensor consider two vectors, $\overrightarrow{A}$ and $\overrightarrow{B}$, and form the combinations

$$T_{ij} = A_i B_j - A_j B_i, \tag{47}$$

which are clearly antisymmetric, and also form a second rank tensor because the components transform like products of components of vectors. In matrix form this is

$$T = \begin{pmatrix} 0 & T_{12} & T_{13} \\ -T_{12} & 0 & T_{23} \\ -T_{13} & -T_{23} & 0 \end{pmatrix}. \tag{48}$$

But, you might say:

> "Isn't this a vector product? This should surely be a vector, but now you tell me that it is an antisymmetric second rank tensor. What is going on?"

In fact, your observation is correct, and, as we shall see in Sec. V, a vector and an antisymmetric second rank tensor are (almost) the same thing. If we write the vector product

$$\overrightarrow{C} = \overrightarrow{A} \times \overrightarrow{B}, \tag{49}$$

then the correspondence is

$$\begin{aligned} C_1 &= T_{23} = -T_{32} \\ C_2 &= T_{31} = -T_{13} \\ C_3 &= T_{12} = -T_{21} \end{aligned} \tag{50}$$

See Sec. V for further discussion of vector products.

IV. CONTRACTIONS

Consider a quantity which transforms like a second rank tensor, A_{ij} say. Then suppose that we set the indices equal and sum to get A_{ii}. How does A_{ii} transform? (Note that the repeated index, i, is summed over.) To answer this question note that

$$\begin{aligned} A'_{ii} &= U_{ij} U_{ik} A_{jk} \\ &= U^T_{ji} U_{ik} A_{jk} \\ &= \left(U^T U\right)_{jk} A_{jk} \\ &= \delta_{jk} A_{jk} && (51) \\ &= A_{jj}, && (52) \end{aligned}$$

where Eq. (51) follows because U is orthogonal. Hence A_{ii} is invariant and so it is a *scalar*. Hence, setting two indices equal and summing has reduced the the rank of the tensor by two, from two to zero (remember, a scalar is a tensor of rank 0). Such a process is called a *contraction*.

Following the same argument one sees the rank of *any tensor* is reduced by two if one sets two indices equal and sums, *e.g.* A_{ijkil} is a third rank tensor (only the unsummed indices, j, k, l, remain). Consequently, the transformation properties of a tensor are determined only by the unsummed indices. This also applies to products of tensors which, as we discussed at the end of the last section, also transform as tensors. Thus, if A, B, C, D and F are tensors (of rank 2, 1, 1, 1 and 4 respectively), then

$$B_i C_i \tag{53}$$
$$\partial B_i / \partial x_i \tag{54}$$
$$A_{ij} B_j$$
$$F_{ijkl} A_{kl}$$
$$F_{ijkl} B_j C_k D_l \tag{55}$$

transform as tensors of rank, 0 (*i.e.* a scalar), 0, 1, 2, and 1 respectively. Note that Eq. (53) is just the scalar product of the vectors $\overrightarrow{B}$ and $\overrightarrow{C}$, and Eq. (54) is a divergence. In the past you always *assumed* that the scalar product and divergence are scalars but probably did not *prove* that they are really invariant on rotating the coordinates.

V. WHY IS A VECTOR PRODUCT A VECTOR?

In the last section we showed why a scalar product really is a scalar (as its name implies). In this section we will prove the analogous result for a vector product. We will deal exclusively with three dimensional vectors, because the vector product is not defined if the vectors are confined to a plane. In Sec. 3, we showed that a second rank antisymmetric tensor looks like a vector product, see Eq. (50). However, we always thought that a vector product is a vector, so how can it also be an antisymmetric second rank tensor? To see that it is indeed both of these things, note that the equation for a vector product, Eq. (49), can be written

$$C_i = \epsilon_{ijk} A_j B_k, \tag{56}$$

where ϵ_{ijk}, call the Levi-Civita symbol, is given by

$$\begin{aligned} &\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \\ &\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \\ &\text{all other } \epsilon_{ijk} = 0. \end{aligned} \tag{57}$$

Clearly ϵ_{ijk} is totally antisymmetric with respect to interchange of any pair of indices.

Because of the effects of contractions of indices, discussed in Sec. IV, C_i will indeed be a vector under rotations if ϵ_{ijk} is an isotropic third rank tensor.[10] This means that if ϵ_{ijk} is given

[10] Note that it is not enough that the notation *suggests* it is a tensor. We have to *prove* that it has the correct transformation properties under rotations.

by Eq. (57), then, in a rotated coordinate system,

$$\epsilon'_{ijk} = U_{il} U_{jm} U_{kn} \epsilon_{lmn} \tag{58}$$

will be equal to ϵ_{ijk}. This is proved in Appendix C.

Consequently, if one forms the antisymmetric second rank tensor

$$\epsilon_{ijk} A_j B_k, \tag{59}$$

then the three components of it (specified by the index i) transform into each other under rotation like a vector. Repeated indices do not contribute to the tensor structure as discussed in Sec. (IV). Hence we have shown that a vector product really is a vector under rotations.

However, a vector product is not quite the same as a vector because it transforms differently under inversion[11], $x_i \to -x_i$. The corresponding transformation matrix is obviously

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{60}$$

which is an orthogonal matrix with determinant -1 (whereas a rotation matrix has determinant $+1$, see Appendix C). Whereas a vector changes sign under inversion, $A_i \to -A_i$, the vector product does not because both A_j and B_k in Eq. (59) change sign. If we wish to make a distinction between the transformation properties of a true vector and a vector product under inversion, we call a true vector a *polar vector* and call a quantity which comes from a vector product a *pseudovector*.[12] As examples, the angular momentum, $\overrightarrow{L} = \overrightarrow{x} \times \overrightarrow{p}$, and the angular velocity, related to the polar vectors $\overrightarrow{v}$ and $\overrightarrow{x}$ by $\overrightarrow{v} = \overrightarrow{\omega} \times \overrightarrow{x}$, are pseudovectors. We say that polar vectors have *odd parity* and pseudovectors have *even parity*.

Similarly, we have to consider behavior of tensors of higher rank under inversion. If $T_{i_1 i_2 \cdots i_n}$ transforms like $x_{i_1} x_{i_2} \cdots x_{i_n}$ under inversion as well as under rotations, *i.e.* if

$$T_{i_1 i_2 \cdots i_n} \to (-1)^n T_{i_1 i_2 \cdots i_n}, \tag{61}$$

under inversion, we say that $T_{i_1 i_2 \cdots i_n}$ is a tensor of rank n, whereas if $T_{i_1 i_2 \cdots i_n}$ changes in the opposite manner, *i.e.*

$$T_{i_1 i_2 \cdots i_n} \to -(-1)^n T_{i_1 i_2 \cdots i_n}, \tag{62}$$

then we say that it is a *pseudotensor*. As an example. the Levi-Civita symbol, ϵ_{ijk} is a third rank *pseudotensor* because, since its elements are constants, it does not change sign under inversion, whereas it would change sign if it were a true tensor. Clearly the product of two pseudotensors is a tensor and the product of a pseudotensor and a tensor is a pseudotensors.

Incidentally, relationships involving vector products can be conveniently derived from the following property of the ϵ_{ijk}:

$$\epsilon_{ijk} \epsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl} \tag{63}$$

(i is summed over). The right hand side is 1 if $j = l$ and $k = m$, is -1 if $j = m$ and $k = l$, and is zero otherwords. By considering the various possibilities one can check that the left hand side takes the same values.

[11] Often called the parity operation in physics.

[12] Also called an axial vector.

As an application of this consider the triple vector product $\overrightarrow{A} \times (\overrightarrow{B} \times \overrightarrow{C})$. Its i-th component is

$$\left[\overrightarrow{A} \times (\overrightarrow{B} \times \overrightarrow{C})\right]_i = \epsilon_{ijk} A_j (\overrightarrow{B} \times \overrightarrow{C})_k = \epsilon_{ijk}\epsilon_{klm} A_j B_l C_m. \tag{64}$$

The value of ϵ_{ijk} is invariant under the "cyclic permutation" $i \to j, j \to k, k \to i$, and so we can write the last expression as

$$\left[\overrightarrow{A} \times (\overrightarrow{B} \times \overrightarrow{C})\right]_i = \epsilon_{kij}\epsilon_{klm} A_j B_l C_m, \tag{65}$$

and use Eq. (63), which gives

$$\left[\overrightarrow{A} \times (\overrightarrow{B} \times \overrightarrow{C})\right]_i = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) A_j B_l C_m = B_i(\overrightarrow{A} \cdot \overrightarrow{C}) - C_i(\overrightarrow{A} \cdot \overrightarrow{B}) \, . \tag{66}$$

Since this is true for all components i we recover the usual result for a triple vector product

$$\overrightarrow{A} \times (\overrightarrow{B} \times \overrightarrow{C}) = \overrightarrow{B}(\overrightarrow{A} \cdot \overrightarrow{C}) - \overrightarrow{C}(\overrightarrow{A} \cdot \overrightarrow{B}) \, . \tag{67}$$

Another result involving vector products which is obtained even more easily from Eq. (63) is

$$(\overrightarrow{A} \times \overrightarrow{B}) \cdot (\overrightarrow{C} \times \overrightarrow{D}) = (\overrightarrow{A} \cdot \overrightarrow{C})(\overrightarrow{B} \cdot \overrightarrow{D}) - (\overrightarrow{A} \cdot \overrightarrow{D})(\overrightarrow{B} \cdot \overrightarrow{C}) \, . \tag{68}$$

VI. TENSOR STRUCTURE OF EQUATIONS

In this section we discuss briefly one of the main reasons why tensors are important in physics. Clearly we want the equations of physics to be valid in any coordinate system. This means that both sides of the equations must have the same tensor structure so they transform in the same way if the coordinate system is rotated. As an example, if the left hand side is a vector then the right hand side must also be a vector. This is illustrated by the equation for angular momentum that we discussed earlier,

$$L_i = I_{ij}\omega_j. \tag{69}$$

Now the angular momentum, L, is a (pseudo) vector, as is the angular velocity, ω_i. As we showed earlier, the moment of inertia, I, is a second rank tensor. Hence, because of the contraction on the index j, the right hand side is a (pseudo) vector, the same as the left hand side. This tells us that Eq. (69) will be true in all rotated frames of reference. Another example is provided by elasticity theory. The stress, σ_{ij}, and strain, e_{ij}, are second rank tensors, and the elastic constant, C_{ijkl}, is a fourth rank tensor. Hence the usual equation of elasticity, which states that the stress is proportional to the strain,

$$\sigma_{ij} = C_{ijkl} e_{kl}, \tag{70}$$

has the same tensor structure on both sides and so will be true in any coordinate system.

To summarize, if we know that the quantities in an equation really are tensors of the form suggested by their indices, one can tell if an equation has the same transformation properties on both sides, and hence is a valid equation, just by *looking* to see if the non-contracted indices are the same. No calculation is required! This is an important use of tensor analysis in physics.

VII. NON-CARTESIAN TENSORS

It is also frequently necessary to consider the transformation properties of quantities under transformations other than rotation. Perhaps the most common example in physics, and the only one we shall discuss here, is the Lorentz transformation in special relativity,

$$\begin{aligned} x' &= \frac{x - vt}{\sqrt{1-(v/c)^2}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t - vx/c^2}{\sqrt{1-(v/c)^2}}. \end{aligned} \tag{71}$$

This describes a transformation between the coordinates in two inertial[13] frames of reference, one of which is moving with velocity v in the x-direction relative to the other. c is, of course, the speed of light. It is more convenient to use the notation

$$\begin{aligned} x^0 &= ct \\ x^1 &= x \\ x^2 &= y \\ x^3 &= z, \end{aligned} \tag{72}$$

where x^μ is called a 4-vector. We will follow standard convention and indicate an index which runs from 0 to 3 by a Greek letter, *e.g.* μ, and an index which just runs over the spatial coordinates (1-3) by a Roman letter, *e.g.* i. The Lorentz transformation, Eq. (71), can be written

$$\begin{aligned} x^{0'} &= \gamma\left(x^0 - \beta x^1\right) \\ x^{1'} &= \gamma\left(x^1 - \beta x^0\right), \end{aligned} \tag{73}$$

neglecting the components which do not change, where

$$\beta = \frac{v}{c}, \qquad \text{and} \quad \gamma = \frac{1}{\sqrt{1-(v/c)^2}}. \tag{74}$$

Note that this can be written as

$$\begin{pmatrix} x^{0'} \\ x^{1'} \end{pmatrix} = U \begin{pmatrix} x^0 \\ x^1 \end{pmatrix}, \tag{75}$$

where

$$U = \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} \tag{76}$$

and θ, often called the rapidity, is given by[14]

$$\tanh\theta = \frac{v}{c}. \tag{77}$$

[13] An inertial frame is one which has no acceleration.

[14] Remember that $1 - \tanh^2\theta = \mathrm{sech}^2\,\theta$.

Eq. (75) is somewhat reminiscent of the rotation matrix that we discussed in earlier sections. However, a significant difference is that the matrix is *not* orthogonal, since, from Eq. (76),

$$U^{-1} = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \neq U^T. \tag{78}$$

Physically, the reason that U is not orthogonal is that $x^\mu x^\mu = \vec{x}^2 + (ct)^2$ is *not* the same in different inertial frames, but rather it is

$$\vec{x}^2 - (ct)^2 \tag{79}$$

which is invariant, because the speed of light is the same in all inertial frames.

To deal with this minus sign we introduce the important concept of the *metric tensor*, $g^{\mu\nu}$, which is defined so that

$$g_{\mu\nu} x^\mu x^\nu \tag{80}$$

is invariant. For the Lorentz transformation, Eq. (79) gives[15]

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{81}$$

We also distinguish between the four vector x^μ, introduced above, which we now call a *contravariant* 4-vector and the *covariant* 4-vector, x_μ, defined by

$$x_\mu = g_{\mu\nu} x^\nu. \tag{82}$$

Note that it is very important to distinguish between upper and lower case indices. Clearly, for a Lorentz transformation, x^μ and x_μ just differ in the sign of the time component,

$$\begin{aligned} x_0 &= -ct \qquad (= -x^0) \\ x_1 &= x \qquad (= x^1) \\ x_2 &= y \qquad (= x^2) \\ x_3 &= z \qquad (= x^3). \end{aligned} \tag{83}$$

Furthermore, from Eq. (80) and (82) we see that the invariant quantity can be expressed as

$$x_\mu x^\mu, \tag{84}$$

i.e. like the scalar product of a contravariant and covariant vector. As shown in Appendix A, a covariant 4-vector transforms with a matrix $V = (U^{-1})^T$, *i.e.*

$$\begin{pmatrix} x_0{}' \\ x_1{}' \end{pmatrix} = V \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \tag{85}$$

[15] g is often defined to be the negative of this. Clearly either sign will work, but the sign in Eq. (81) seems more natural to me.

where, from Eq. (78),

$$V = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}. \tag{86}$$

One can define other contravariant 4-vectors, A^μ say, as quantities which transform in the same way as x^μ, and other covariant 4-vectors, A_μ, which transform like x_μ. We show in Appendix B that the gradient of a scalar function with respect to the components of a contravariant vector is a covariant vector and vice-versa. Hence frequently

$$\frac{\partial\phi}{\partial x_\mu} \quad \text{is written as} \quad \partial^\mu\phi, \tag{87}$$

because the latter form directly indicates the tensor structure. Note that the index μ is superscripted in the expression on the right but subscripted in the expression on the left. Similarly

$$\frac{\partial\phi}{\partial x^\mu} \quad \text{is written as} \quad \partial_\mu\phi. \tag{88}$$

One can also define higher order contravariant tensors, which can transform like products of contravariant 4-vectors. These have upper case indices like $C^{\mu\nu\lambda}$. One can also define covariant tensors which have lower case indices like $C_{\mu\nu\lambda}$. One can also define tensors with mixed covariant-contravariant transformation properties. These have some upper case and some lower case indices, *e.g.* $C^{\lambda}_{\mu\nu}$.

Because $x^\mu x_\mu$ is invariant, contractions are obtained by equating a covariant index and a contravariant index and summing over it. Some examples are

$$\begin{aligned} & A_\mu B^\mu \\ & \partial_\mu B^\mu \equiv \frac{\partial B^\mu}{\partial x^\mu} \\ & C_{\mu\nu} B^\nu \\ & D^{\mu\nu}_{\lambda\sigma} C^{\lambda\sigma}, \end{aligned} \tag{89}$$

which give, respectively, a scalar, a scalar, a covariant 4-vector and a contravariant second rank tensor.

You might say that we have generated quite a lot of extra formalism, such as the metric tensor and two types of vectors, just to account for the minus sign in Eq. (79) when we deal with special relativity, and this is true. Nonetheless, though not essential for special relativity, it is quite convenient to use tensor formalism for this topic because the method is so elegant. Furthermore, in general relativity one has a *curved* space-time, as a result of which the metric tensor does not have the simple form in Eq. (81), and is a function of position. The situation is much more complicated, so tensor analysis *is* essential for general relativity.

APPENDIX A: ORTHOGONALITY OF THE ROTATION MATRIX

In this section we prove that a rotation matrix is orthogonal. We also use a similar argument to show how a covariant vector, defined in Sec. VII, transforms.

We start by noting that, under a rotation, the length of a vector is preserved so

$$x_i x_i = x'_j x'_j, \tag{A1}$$

(remember i and j are summed over) or equivalently

$$x_i x_i = U_{jk} U_{jl} x_k x_l. \tag{A2}$$

This implies that

$$U_{jk} U_{jl} = \delta_{kl}, \tag{A3}$$

where δ_{kl}, the Kronecker delta function, is 1 if $k = l$ and 0 otherwise. Eq. (A3) can be written

$$U^T_{kj} U_{jl} = \delta_{kl}, \tag{A4}$$

(where T denotes the transpose), which can be expressed in matrix notation as

$$U^T U = I, \tag{A5}$$

(where I is the identity matrix). Eq. (A5) is the definition of an orthogonal matrix.[16]

A similar argument shows how a covariant vector transforms. If a contravariant vector, x^μ say, transforms like

$$x^{\mu\prime} = U_{\mu\nu} x^\nu, \tag{A6}$$

then a covariant vector, x_μ, transforms according to

$$x'_\mu = V_{\mu\nu} x_\nu, \tag{A7}$$

say. The goal is to determine the matrix V. According to Eq. (84) $x^\mu x_\mu$ is invariant, and so, from Eqs. (A6) and (A7),

$$\begin{aligned} x^\nu x_\nu &= x^{\mu\prime} x'_\mu \\ &= U_{\mu\lambda} V_{\mu\sigma} x^\lambda x_\sigma. \end{aligned} \tag{A8}$$

which implies, as in Eq. (A5),

$$U^T V = I, \tag{A9}$$

Hence we obtain

$$V = \left(U^{-1}\right)^T. \tag{A10}$$

Note that for cartesian tensors, U is a rotation matrix, which is orthogonal, and so $V = U$. Hence it is not necessary in this (important) case to distinguish between covariant and contravariant vectors.

---

[16] Note that another, equivalent, definition of an orthogonal matrix is that the columns (and also the rows) form orthonormal vectors. This follows directly from Eq. (A3).

APPENDIX B: TRANSFORMATION OF DERIVATIVES

In this section we discuss the tensorial properties of derivatives. First we discuss the case of transformation under rotations, *i.e.* we consider cartesian tensors.

Consider a scalar function ϕ. The derivatives will transform according to

$$\frac{\partial \phi}{\partial x'_i} = V_{ij} \frac{\partial \phi}{\partial x_j}, \tag{B1}$$

where V is a matrix of coefficients which we want to determine. If $V = U$, then the gradient is indeed a vector. To show that this is the case, start by noting that

$$\frac{\partial \phi}{\partial x'_i} = \frac{\partial \phi}{\partial x_j} \frac{\partial x_j}{\partial x'_i}. \tag{B2}$$

Furthermore,

$$\overrightarrow{x'} = U \overrightarrow{x}, \tag{B3}$$

so

$$\overrightarrow{x} = U^{-1} \overrightarrow{x'}, \tag{B4}$$

and hence

$$\frac{\partial x_j}{\partial x'_i} = \left(U^{-1}\right)_{ji}. \tag{B5}$$

Consequently, Eq. (B2) can be written

$$\frac{\partial \phi}{\partial x'_i} = \frac{\partial \phi}{\partial x_j} \left(U^{-1}\right)_{ji}, \tag{B6}$$

or

$$\frac{\partial \phi}{\partial x'_i} = \left(U^{-1}\right)^T_{ij} \frac{\partial \phi}{\partial x_j}, \tag{B7}$$

and comparing with Eq. (B1), we see that

$$V = \left(U^{-1}\right)^T \tag{B8}$$

$$= U, \tag{B9}$$

where the last equality follows because U is an orthogonal matrix and so the inverse is equal to the transpose. Hence, when we are referring to transformations under ordinary rotations, the gradient of a scalar field is a vector.

When the transformation is not orthogonal, as for example in the Lorentz transformation, then only the last equality no longer applies, so Eq. (B8), which is still valid, shows that derivatives transform with the matrix

$$V = \left(U^{-1}\right)^T. \tag{B10}$$

We show in Appendix A that if a contravariant vector, x^μ, transforms with a matrix U, then a covariant vector transforms with a matrix $\left(U^{-1}\right)^T$. Hence Eqs. (B1) and (B10) shows that the gradient of a scalar function differentiated with respect to the contravariant vector x^μ is a covariant vector. The converse is also easy to prove.

APPENDIX C: INVARIANCE OF LEVI-CIVITA SYMBOL

In this appendix we sketch a proof that ϵ_{ijk} is an isotropic third rank tensor. Let us consider the case of $i = 1, j = 2, k = 3$. From Eq. (58) we have

$$\begin{aligned}
\epsilon'_{123} &= U_{1l}U_{2m}U_{3n}\epsilon_{lmn} && \text{(C1)}\\
&= \quad U_{11}U_{22}U_{33} \qquad (l=1, m=2, n=3) \\
&\quad - U_{11}U_{23}U_{32} \qquad (l=1, m=3, n=2) \\
&\quad - U_{12}U_{21}U_{33} \qquad (l=2, m=1, n=3) \\
&\quad + U_{12}U_{23}U_{31} \qquad (l=2, m=3, n=1) \\
&\quad + U_{13}U_{21}U_{32} \qquad (l=3, m=1, n=2) \\
&\quad - U_{13}U_{22}U_{31} \qquad (l=3, m=2, n=1) && \text{(C2)}\\
&= \begin{vmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{vmatrix} && \text{(C3)}\\
&= \det U && \text{(C4)}\\
&= 1 && \text{(C5)}\\
&= \epsilon_{123}. && \text{(C6)}
\end{aligned}$$

To verify Eq. (C3) just expand out the determinant and check that is the same as Eq. (C2). Eq. (C5) follows because U is orthogonal, so $UU^T = 1$, and since $\det AB = \det A \det B$, we have $\det U \det U^T = 1$. Now clearly $\det U = \det U^T$ so $(\det U)^2 = 1$ and hence

$$\det U = \pm 1, \tag{C7}$$

for a real orthogonal matrix. A rotation corresponds to the plus sign, since, for example, a rotation by θ about the z axis is represented by

$$U = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{C8}$$

whose determinant is $\cos^2\theta + \sin^2\theta = 1$. An example of an orthogonal transformation which which has determinant -1 is inversion, $x'_i = -x_i$.

One can repeat the arguments which led to Eq. (C6) for other values of i, j and k all distinct. In each case one ends up with $\pm\det U = \pm 1$, with, in all cases, the sign the same as that of ϵ_{ijk}. If two or more of the indices i, j and k are equal, then ϵ'_{ijk} is clearly zero. To see this, consider for example

$$\epsilon'_{11k} = U_{1l}U_{1m}U_{kn}\epsilon_{lmn}. \tag{C9}$$

Because ϵ_{lmn} is totally antisymmetric, contributions from pairs of terms where l and m are interchanged cancel, and hence $\epsilon'_{11k} = 0$. We have therefore proved that, for all elements,

$$\epsilon'_{ijk} = \epsilon_{ijk}. \tag{C10}$$

(under rotations). However, under inversion (for which $\det U = -1$), ϵ_{ijk} would change sign if it were a genuine tensor, whereas its elements are constant. Hence the Levi-Civita symbol is an isotropic, completely antisymmetric, third rank *pseudo*tensor.