An Introduction to Tensors

A.C. Norman
anorman@bishopheber.cheshire.sch.uk

November 2010

Abstract

As we shall see, not all things in physics are created equal; in order to properly describe the world around us, we physicists have to resort to more than one type of mathematical object to describe physical situations. You should already be familiar with scalar and vector quantities, so we shall begin with some revision of their mathematical properties, before introducing and investigating tensors of higher order.

1 Scalars

Quantities such as mass, energy, charge, length or temperature are known as *scalars*. As you will already know, they each have a magnitude, but no direction, and this means that they may be treated as simple numbers algebraically.

We shall discover that it is of some importance that scalars remain unchanged or *invariant* under a rotation of our coordinate frame. A scalar S will in a rotated system be equal to S', where

$$S' = S. \tag{1}$$

Simple examples of scalars from mathematics are the length of a vector, and the dot product of two vectors. It follows that any physically important quantities formed from a scalar product are also scalars, and perhaps the most immediate of these is energy, either as potential energy or as an energy density (e.g. $\mathbf{F} \cdot \mathrm{d}\mathbf{r}$, $e\mathbf{E} \cdot \mathrm{d}\mathbf{r}$, $\mathbf{D} \cdot \mathbf{E}$, $\mathbf{B} \cdot \mathbf{H}$, $\boldsymbol{\mu} \cdot \mathbf{B}$), but others, such as the angle between two directed quantities, are also important.

2 Vectors

Vectors have both a size and a direction, and you will be familiar with this from your studies.

Let's first look at a type of object which is *not* a vector. We could describe rotations of rigid bodies in the following way. A certain rotation is represented by an arrow, with its direction along the axis of rotation in a sense given by the right hand rule, and length given by the rotation angle in radians. Apparently then, a rotation is a vector according to the size and direction definition. However, you can show that the arrows we have associated with rotations are not vectors[1], since they do not behave like vectors. Take a book, and rotate it by $90°$ around the x axis, then $90°$ around the y axis. Starting from exactly the same orientation, repeat these operations, but this time about the y axis and then the x axis. The final orientations of the book are different, but we know that, for vectors, the order in which they are added together does not matter (we say that vector addition is commutative).

We are clearly in need of a better definition of a vector, since simply having a size and a direction is not enough for an object to be a vector. Let us now consider the idea of a vector as a set of components – three in a three dimensional space. To talk about components, we need a coordinate system, and we see that we can pick any from an infinite variety (even if we stick with Cartesian axes, there are an uncountable infinity of rotated axes). A vector consists of thee components *in each coordinate system*. We can find the components in another system, of course, by taking projections, meaning that the new components will be definite combinations of the old components. This allows us to decide whether a physical quantity is really a vector or not.

Mathematically, we could *define* any transformation laws between coordinate systems, so long as they allow components in one system to be made into a set of components for all other systems. However, when dealing with physical entities, we are not free to define its components in various coordinate systems, since they are defined by physical fact. We again see now why our rotation arrows were not vectors: if we treat such an arrow as a vector and take components of it, these will not represent rotations which can be combined to give the original rotation. The vector looking superficially like the arrow we defined is not a correct mathematical description of the physical entity (rotation) which we are trying to describe.

It may seem obvious that all relations between physical quantities, that are the quantitative descriptions of physical process, must be independent of the measuring scale and the frame of reference used. However, we can turn this argument around: since physical results must obey those laws of symmetry meaning that they will be independent of the choice of coordinate system, what does this imply about the nature of quantities involved in the description of physical processes?

The fact that a physical relationship can be expressed as a vector equation assures us that the relationship is unchanged by a mere rotation of the coordinate system. That is the reason why vectors are so useful in physics.

We shall now investigate how components in one coordinate system are related to those in another coordinate system for vectors. We start in a coordinate system with basis vectors along the x, y and z axes. In many books, these are written as $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, but here I am going to use $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ so

[1]Of course, we **do** define rotations this way, but they are not 'honest' vectors in the usual sense. Examples of such vectors are the torque $\boldsymbol{\tau}$, angular momentum $\mathbf{L}$, angular velocity $\boldsymbol{\omega}$ and magnetic field $\mathbf{B}$. They are known as *axial vectors* or *pseudovectors*. Ordinary vectors, such as position $\mathbf{r}$, force $\mathbf{F}$, momentum $\mathbf{p}$, electric field $\mathbf{E}$, and so forth, are known as *polar vectors*.

that results can be easily generalized to any number of dimensions. We are going to examine the vector $\mathbf{x}$, which has components x_1, x_2, x_3:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \sum_i x_i\mathbf{e}_i, \tag{2}$$

where in the final form the sum is over the dimensions of the space.

We now introduce a new basis $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$, which is related to the old one by a rotation, keeping the origin fixed. Each of the vectors $\mathbf{e}'_i$ may be expressed in terms of the original coordinate set as follows:

$$\begin{aligned}
\mathbf{e}'_1 &= S_{11}\mathbf{e}_1 + S_{21}\mathbf{e}_2 + S_{31}\mathbf{e}_3 \\
\mathbf{e}'_2 &= S_{12}\mathbf{e}_1 + S_{22}\mathbf{e}_2 + S_{32}\mathbf{e}_3 \\
\mathbf{e}'_3 &= S_{13}\mathbf{e}_1 + S_{23}\mathbf{e}_2 + S_{33}\mathbf{e}_3
\end{aligned} \tag{3}$$

The vector $\mathbf{x}$ can be related to either of the coordinate systems by means of components

$$\mathbf{x} = x'_1\mathbf{e}'_1 + x'_2\mathbf{e}'_2 + x'_3\mathbf{e}'_3 = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

Taking the dot product of this equation with $\mathbf{e}'_1$, we get

$$\begin{aligned}
\mathbf{x}\cdot\mathbf{e}'_1 = x'_1 =& x_1\mathbf{e}_1\cdot\mathbf{e}'_1 + x_2\mathbf{e}_2\cdot\mathbf{e}'_1 + x_3\mathbf{e}_3\cdot\mathbf{e}'_1 \\
=& x_1\mathbf{e}_1\cdot(S_{11}\mathbf{e}_1 + S_{21}\mathbf{e}_2 + S_{31}\mathbf{e}_3) + \\
& x_2\mathbf{e}_2\cdot(S_{11}\mathbf{e}_1 + S_{21}\mathbf{e}_2 + S_{31}\mathbf{e}_3) + \\
& x_3\mathbf{e}_3\cdot(S_{11}\mathbf{e}_1 + S_{21}\mathbf{e}_2 + S_{31}\mathbf{e}_3) \\
x'_1 =& S_{11}x_1 + S_{21}x_2 + S_{31}x_3.
\end{aligned}$$

Similarly, by dotting $\mathbf{x}$ into $\mathbf{e}'_1$ and $\mathbf{e}'_3$, we get

$$x'_2 = S_{12}x_1 + S_{22}x_2 + S_{32}x_3,$$

$$x'_3 = S_{13}x_1 + S_{23}x_2 + S_{33}x_3,$$

which are known as the transformation equations from the coordinate system $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$. They may be more concisely written in matrix notation. If we denote the matrix with elements S_{ij} as S, then the new basis is related to the old one by

$$\mathbf{e}'_j = \sum_i S_{ij}\mathbf{e}_i, \tag{4}$$

and the components x'_i and x_i are related by

$$x'_i = \sum_j (\mathsf{S}^{-1})_{ij}x_j. \tag{5}$$

In the special case that the transformation is a rotation, the transformation matrix S is orthogonal and so the inverse matrix S^{-1} is the same as the transpose matrix S^{T}, so

$$x'_i = \sum_j (\mathsf{S}^{\mathrm{T}})_{ij}x_j = \sum_j \mathsf{S}_{ji}x_j. \tag{6}$$

3 Tensors

Tensors are simply a generalization of scalars and vectors. In fact, scalars and vectors are themselves tensors of order (or rank[2]) zero and one respectively. In three-dimensional space a scalar has one (or 3^0) component, and a vector has 3 (or 3^1) components; in general, a tensor of order n in D dimensions has D^n components. Let us look at an example of where we might encounter a second-order tensor, having nine components, in a physical application.

Think of a beam carrying a load. Stresses and strains are set up in the material of the beam, which enable it to bear the load. If you imagine cutting the beam in two by a plane perpendicular to its length, you will realize that there is a force per unit area on the material on one side of the cut from the material on the other side (this is what is keeping the beam from snapping; there is, of course, and equal and opposite pressure on the other side of the cut). This force is a vector, and so it has components σ_{xx}, σ_{xy}, σ_{xz}, where the first subscript x emphasizes that this is a force across a plane perpendicular to the x direction. Similarly, if we were to make a cut perpendicular to the y direction, we should find a force per unit area across this plane with components σ_{yx}, σ_{yy}, σ_{yz}, and finally across a plane perpendicular to the z direction there is a force per unit area with components σ_{zx}, σ_{zy}, σ_{zz}. At some point in the material, we have a set of nine quantities, each a component of a force acting on one of three orthogonal planes, which could be displayed as a matrix:

$$\begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}. \tag{7}$$

This is a second-order tensor known as the stress tensor. The components σ_{xx}, σ_{yy} and σ_{zz} are pressures or tensions, and the others are shear forces (per unit area). For example, σ_{zx} is a force per unit area in the x direction acting across a plane perpendicular to the z direction.

We therefore expect a second-order tensor to have nine coordinates (in three dimensions) in every rectangular coordinate system. Let us construct a second-order tensor to find out what happens to its components when transforming from one coordinate system to another. Our very simple example is formed from the components of two vectors, U and V:

$$\begin{matrix} U_1V_1 & U_1V_2 & U_1V_3 \\ U_2V_1 & U_2V_2 & U_2V_3 \\ U_3V_1 & U_3V_2 & U_3V_3 \end{matrix} \tag{8}$$

These are the components of a second-order tensor which we shall denote by **UV** (*note*: no dot or cross). Since **U** and **V** are vectors, their components in a rotated coordinate system are, by (5):

$$U'_k = \sum_i a_{ki} U_i, \qquad V'_l = \sum_j a_{lj} V_j.$$

Hence the components of the second-order tensor **UV** are

$$U'_k V'_l = \sum_i a_{ki} U_i \sum_j a_{lj} V_j = \sum_{i,j} a_{ki} a_{lj} U_i V_j, \tag{9}$$

[2]In some books you will see *order*, and in others, *rank*; the two terms are equivalent.

which allows us to write the form for a general second-order tensor T_{ij} transforming to a coordinate system in which it is written as T'_{kl}, given by

$$T'_{kl} = \sum_{i,j} a_{ki} a_{lj} T_{ij}. \tag{10}$$

Equation (10) generalizes immediately. For example, a fourth-order Cartesian tensor is defined by 3^4 or 81 components T_{ijkl}, in every rectangular coordinate system, which transform to a rotated coordinate system by the equation

$$T'_{\alpha\beta\gamma\delta} = \sum_{ijkl} a_{\alpha i} a_{\beta j} a_{\gamma k} a_{\delta l} T_{ijkl}. \tag{11}$$

3.1 Einstein Summation Convention

As a simplification of notation, it is customary to omit the summation signs in equations like (2), (4), (5) and (6), and simply understand that summation is carried out in such expressions over any *repeated* subscript. Thus

$$a_i a_i \text{ or } a_j a_j \text{ or } a_f a_f \text{ means } a_1^2 + a_2^2 + a_3^2;$$

$$a_{ij} a_{jk} \text{ means } a_{i1} a_{1k} + a_{i2} a_{2k} + a_{i3} a_{3k}.$$

The repeated index which is summed over is know as a *dummy* index. Since it disappears, like a variable of integration, it doesn't matter what letter is used for it.

4 Uses of tensors

In this section, some physical applications of tensors are given, albeit briefly. First-order tensors are already familiar as vectors, and so we shall concentrate on second-order tensors, starting with an example taken from mechanics.

4.1 Rigid body rotation

If a rigid body is rotating about a fixed axis, then $\mathbf{L} = I\boldsymbol{\omega}$ is a correct vector equation, where $\mathbf{L}$ is the angular momentum, I is the moment of inertia about the rotation axis, and $\boldsymbol{\omega}$ is the angular velocity. As indicated by the equation, $\mathbf{L}$ and $\boldsymbol{\omega}$ are parallel vectors, and I is a scalar.

In general, where the rotation axis is not fixed, the angular velocity and the angular momentum are not parallel. If the equation $\mathbf{L} = I\boldsymbol{\omega}$ is to be true, I cannot be a scalar; it must instead be some quantity which, multiplied by $\boldsymbol{\omega}$, gives a vector in a different direction. In fact, in general, I is a second-order tensor. Written out in full Cartesians, the inertia tensor for a continuous body would have the form

$$\mathsf{I} = [I_{ij}] = \begin{pmatrix} \int (y^2+z^2)\rho\,\mathrm{d}V & -\int xy\rho\,\mathrm{d}V & -\int xz\rho\,\mathrm{d}V \\ -\int xy\rho\,\mathrm{d}V & \int (z^2+x^2)\rho\,\mathrm{d}V & \int yz\rho\,\mathrm{d}V \\ \int xz\rho\,\mathrm{d}V & \int yz\rho\,\mathrm{d}V & \int (x^2+y^2)\rho\,\mathrm{d}V \end{pmatrix}, \tag{12}$$

where $\rho = \rho(x, y, z)$ is the mass distribution and $\mathrm{d}V$ stands for $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ (the integrals are to be taken over the whole body). The diagonal elements of this tensor are called the *moments of inertia* and the off-diagonal elements without the minus signs are called the *products of inertia*.

4.2 Polarization of light

In an isotropic medium the electric polarization $\mathbf{P}$ is a vector parallel to the electric field $\mathbf{E}$, that is, $\mathbf{P} = \chi\mathbf{E}$, where χ is a constant. For anisotropic materials such as crystals, this may no longer be true, and again we need a mathematical quantity which, and again we need a mathematical object which gives a vector in a different direction when multiplied by a vector.

A similar equation relates the magnetic susceptibility to the magnetic field, and the electrical conductivity to the electric field. Usually these are found to be symmetric second-order tensors.

4.3 Elasticity

In the general case, not only are the stress and strain not parallel, they are not even vectors, but are themselves second-order tensors, and the quantity which replaces Y is a fourth-order tensor.

4.4 Relativity

As we have already pointed out, tensors readily generalize to spaces of higher dimensionality. The most notable example of this in physical theory are the more abstract four dimensional space-time of special and general relativity.

References

BOAS, M.L. *Mathematical Methods in the Physical Sciences*, 2nd Edtn., 1983, pp. 435–455

FEYNMAN, R.P *Lectures on Physics*, 1963, I-11, I-20.

RILEY, HOBSON & BENCE *Mathematical Methods for Physics and Engineering*, 2nd Edtn., 2002, pp. 776–833.

SMITH, L.P. *Mathematical Methods for Scientists and Engineers*, 1953, pp. 300–308.