DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY, MADRAS

PH350 Classical Physics **Handout 1** **8.8.2009**

1 Scalars, Vectors and Tensors

In physics, we are interested in obtaining laws (in the form of mathematical equations) which govern the behaviour of different systems. These laws should hold at different places as well as times. They should also be independent of the observer testing them. We thus infer that the physical laws and values of experimentally measurable quantities must be frame independent, i.e., independent of the choice of location of the origin of coordinates and the coordinate axes that we choose. In order to check this, we need to be able to convert results in one frame to equivalent results in another frame.

Suppose we are able to write the physical laws, i.e., the mathematical equations which describe them, in terms of objects that transform in a definite manner under a change of frame, the rule for conversion is manifest. Examples of such objects are scalars, vectors and more generally, tensors. When the equations are written in terms of these objects, we say that the equations are **form invariant**. For instance, an equation such as Newton's second law, $\vec{F} = m\vec{a}$, has vectors on both sides of the equation. So while two different observers have different numerical values for each component of their vectors, they will both agree that the equation holds.

However, the elementary or high-school definition of a vector ("a quantity with both magnitude and direction") is rather unsatisfactory. In this unit, we shall first provide a better definition of a vector through its properties under a change in coordinate axes (change in basis vectors).

1.1 What is a vector?

The **standard** example of a vector in three dimensions is the position vector

$$\vec{r} = x\hat{e}_x + y\hat{e}_y + z\hat{e}_z \quad .$$

The three numbers (x, y, z) specify the vector fully. However, their actual numerical values depend on the particular coordinate axes chosen: In other words, a set $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ of mutually perpendicular unit vectors is chosen, and the values (x, y, z) are with reference to this set[1]. This choice is not unique. A different orthonormal basis will give rise to a different set of numbers, say, (x', y', z') associated with the same position vector. How are the two sets related? It is useful to relabel the coordinates as

$$x_1 = x \ , \ x_2 = y \ , \ x_3 = z \quad .$$

[1] This is called an orthonormal basis. Each vector has unit magnitude, and they are mutually orthogonal (perpendicular): $\hat{e}_x \cdot \hat{e}_y = 0$ etc.

The relation between the two sets of coordinates can be written as

$$x'_i = \sum_{j=1}^{3} R_{ij}\ x_j \quad , \tag{1}$$

where the nine numbers R_{ij} are the coefficients showing how the two sets of axes are related. This can be rewritten in matrix form if we represent the coordinates as a column vector and R_{ij} as the entries in a matrix R with i as the row label and j as the column label.

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad . \tag{2}$$

We can show that the matrix R satisfies the conditions (R^{T} is the transpose of the matrix R)

$$R^{\mathrm{T}} = R^{-1} \quad \text{or} \quad R^{\mathrm{T}}\ R = R\ R^{\mathrm{T}} = I \tag{3}$$

$$\text{and}\ \det R \ = \ 1 \quad , \tag{4}$$

where I is the (3×3) unit matrix. The first condition ensures that the magnitude of vectors remain unchanged in all bases[2]. The second condition preserves orientation, i.e., the basis vectors always form a right-handed triad. Such matrices are called **rotation matrices** because two different choices of basis vectors are related by rotations. For instance, consider the case when two bases have one common basis vector, $\hat{e}'_z = \hat{e}_z$, i.e., they have the same z-axis. The coordinates are related by

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{5}$$

i.e., the two x- and y-axes are related by a rotation by an angle ϕ about the z-axis.

We can now generalise this to arbitrary rotations. Arbitrary rotations are specified by an axis of rotation $\hat{n}$ and an angle of rotation ϕ about the direction of $\hat{n}$. Let us denote the corresponding rotation matrix by $R(\hat{n}, \phi)$. Thus, the rotation matrix given above corresponds to $R(\hat{e}_z, \phi)$. One can give an explicit form for the general rotation matrix but we shall not do so now.

We can now provide the definition of a vector. **"A vector is a set of three quantities that transform, under rotations of the coordinate axes, exactly as the set of coordinates itself transform."**. We usually represent the three quantities as a column vector, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Under the change of coordinates given by Eqn.(1), the components transform as

$$v'_i = \sum_{j=1}^{3} R_{ij}\ v_j \quad . \tag{6}$$

[2]Actually it also ensures that the new basis vectors are unit vectors and orthogonal to each other.

Exercise Given a vector $\vec{v}$, find a rotation matrix R (or equivalently, a change of basis) such that $\vec{v}' = (0, 0, v)^{\mathrm{T}}$. Is R unique? Explain.

1.2 The Einstein summation convention

This convention is very useful in writing in a compact and uncluttered form, formulae involving several summations. The convention is simply that any index that occurs twice is summed over. The summation always runs from 1 to 3 since we are in three dimensions. Further, no index can occur more than twice in any equation. Thus, we can rewrite Eqn. (6) as

$$v'_i = R_{ij}\ v_j \tag{7}$$

where the summation of j is implicitly assumed under the Einstein summation convention and i is a **free** (i.e., unsummed) **index**. The index that is summed over is referred to as a **dummy index**. Care should be taken to see that different letters are used to indicate different summations. **Further, a simple check is to see that the same free indices occur on either side of any equation**.

1.3 The dot-product: Scalars from vectors

A **scalar** is defined as an object that is unchanged or **invariant** under change of basis/rotations.

Given two vectors $\vec{v} = (v_1, v_2, v_3)^{\mathrm{T}}$ and $\vec{w} = (w_1, w_2, w_3)^{\mathrm{T}}$, we can construct a **scalar** from them. It is called the **dot product**, and is defined by

$$\vec{v} \cdot \vec{w} \equiv v_i\ \delta_{ij}\ w_j \quad , \tag{8}$$

where

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

is the **Kronecker delta**. When it is written as a 3×3 matrix, it is nothing but the identity matrix I. Thus, it is not hard to see that

$$\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v} = v_i w_i \tag{9}$$

(Note the summation over i as per our convention.) The invariance of the dot-product follows from the condition $R^{\mathrm{T}}R = I$ satisfied by rotation matrices. Explicitly, this condition is

$$R^{\mathrm{T}}R = I \rightarrow R_{ji}R_{jk} = \delta_{ik}\ .$$

Here i and k are free indices and the dummy index j is summed over from 1 to 3. The following steps show how the dot product is unchanged when going from one basis to another

$$v'_j w'_j = R_{ji}v_i R_{jk}w_k = R_{ji}R_{jk}v_i w_k = \delta_{ik}v_i w_k = v_i w_i$$

The dot-product when we choose $\vec{v} = \vec{w}$ gives the square of the magnitude of a vector. That is

$$|\vec{v}|^2 = v_i v_i = v_1^2 + v_2^2 + v_3^2 \quad , \tag{10}$$

which is clearly positive definite. Note that $|\vec{v}| = 0$ implies $\vec{v} = 0$. For any two vectors $\vec{u}$ and $\vec{v}$, we can show that

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos\theta \quad ,$$

where θ is the angle between the two vectors $\vec{u}$ and $\vec{v}$. Recall that in elementary vector analysis this formula was used to *define* the scalar product. We now see that it follows as a consequence of the definition $\vec{u} \cdot \vec{v} = u_i v_i$.

1.4 The cross-product: New vectors from old

Given two vectors $\vec{v}$ and $\vec{w}$, one can construct a new vector $\vec{u}$ by the **cross product** defined below: the components of the cross product are

$$\begin{aligned} u_1 &= v_2 w_3 - v_3 w_2 \\ u_2 &= v_3 w_1 - v_1 w_3 \\ u_3 &= v_1 w_2 - v_2 w_1 \quad , \end{aligned} \tag{11}$$

written symbolically as $\vec{u} = \vec{v} \times \vec{w}$. Eqs. (12) can be written in compact form as

$$u_i = \epsilon_{ijk}\ v_j\ w_k \equiv (\vec{v} \times \vec{w})_i \tag{12}$$

where ϵ_{ijk} is the **Levi-Civita tensor** (also called the **permutation symbol**) defined as:

$$\begin{aligned} \epsilon_{123} &= 1 \\ \epsilon_{ijk} &= -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji} \end{aligned}$$

The second condition implies that the Levi-Civita tensor is totally antisymmetric and is non-vanishing if and only if all three indices are distinct. Thus, $\epsilon_{123} = \epsilon 231 = \epsilon_{312} = 1$ and $\epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1$. All other components are vanishing.

For any two vectors $\vec{u}$ and $\vec{v}$, one can show that

$$|\vec{u} \times \vec{v}| = |\vec{u}||\vec{v}| \sin\theta \quad ,$$

where θ is the angle between the two vectors $\vec{u}$ and $\vec{v}$. Clearly, $\vec{u} \times \vec{u} = 0$.

1.5 More complicated objects: Tensors

As mentioned earlier, scalars and vectors are not the only kinds of objects that one encounters in physical situations. A simple example of a more complicated object - a **tensor**, is given by $A_{ij} = v_i w_j$, where $\vec{v}$ and $\vec{w}$ are any two vectors. Under a change of basis, one can see that

$$A'_{ij} = v'_i w'_j = R_{ik}\ R_{jl} v_k w_l = R_{ik}\ R_{jl} A_{kl} \quad .$$

A tensor like A_{ij} can be represented as a 3×3 matrix A. Then, the above transformation law can be written as

$$A' = R^{\mathrm{T}} \; A \; R \tag{13}$$

Objects which transform in this fashion are called **tensors of rank two** (because its components are specified by *two* indices i and j). Not all tensors of rank two can be written as $v_i w_j$ and thus one can use the above equation as the defining condition. An example of such an object is given by the moment of inertia of a rigid body. Other examples include the stress and strain tensors, the dielectric tensor and so on.

Clearly, a tensor of second rank is the simplest generalisation of a vector. Generalising further, tensors of rank r are objects which have r-indices. Under a rotation (given by the rotation matrix R_{ij}), a tensor of rank r, $T_{i_1 i_2 \cdots i_r}$ transforms as

$$T'_{i_1 i_2 \cdots i_r} = R_{i_1 j_1} \; R_{i_2 j_2} \cdots R_{i_r j_r} \; T_{j_1 j_2 \ldots j_r} \; . \tag{14}$$

Thus, **scalars are tensors of rank zero** and **vectors are tensors of rank one**. The Levi-Civita tensor has rank three and the Kronecker delta has rank two. The elastic constants that relate stress and strain in linear materials is a tensor of rank four.

Exercise Using the transformation law for a tensor of rank two, show that the Kronecker delta is invariant under rotations of the coordinate axes.

Exercise Using the transformation law for a tensor of rank three and the antisymmetry of the Levi-Civita tensor, show that

$$\epsilon'_{ijk} = (\det R) \; \epsilon_{ijk} \quad .$$

Thus, the Levi-Civita tensor is invariant under rotations since $\det R = 1$.
Comment: Under arbitrary rotations, the Kronecker delta and the Levi-Civita tensor are the only invariant ('isotropic') tensors. Any higher rank tensor that are isotropic can always be written in terms of these two basic tensor!

1.6 An important identity

As we just saw, the Kronecker delta and Levi-Civita tensor are invariant under change of basis. They are related by the following important identity which enables one to derive just about all vector identities:

$$\epsilon_{ijk}\epsilon_{ilm} = (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) \tag{15}$$

Here we must remember that the index i is repeated, and so a summation over i from $i = 1$ to $i = 3$ is implicit. The most straightforward proof is to explicitly verify that this is true for all values of the free indices (j, k, l, m). This is left to the student. However, some simple checks are: both the LHS and RHS are antisymmetric under the exchanges: $j \leftrightarrow k$ and $k \leftrightarrow l$.

As an application of the identity (15), consider the multiple cross-product: $\vec{u} \times (\vec{v} \times \vec{w})$. The i-th component of this is

$$(\vec{u} \times (\vec{v} \times \vec{w}))_i = \epsilon_{ijk} u_j (\vec{v} \times \vec{w})_k = \epsilon_{ijk} u_j \epsilon_{klm} v_l w_m = \epsilon_{kij} \epsilon_{klm} u_j v_l w_m$$

Using the identity (15), we obtain

$$(\vec{u} \times (\vec{v} \times \vec{w}))_i = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, u_j v_l w_m = u_m v_i w_m - u_j v_j w_i$$

which can be rewritten as

$$(\vec{u} \times (\vec{v} \times \vec{w})) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w} \tag{16}$$

One can further show that

$$(\vec{u} \times \vec{v}) \times \vec{w} \neq \vec{u} \times (\vec{v} \times \vec{w}) \quad ,$$

i.e., the order of the cross-product is important.

1.7 Reflections and Parity

We saw that rotations are given by matrices that satisfy $R^{\mathrm{T}} = R^{-1}$ and $\det R = 1$. Are there transformations which satisfy only the first of these two conditions? Taking the determinant on both sides, we find that $\det R^2 = 1$, which suggests that one consider transformations such that $\det R = -1$. It is not hard to see that reflection about any plane is such an operation. Under reflections about the xy-plane, one obtains

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ -z \end{pmatrix} \quad .$$

The matrix $R = \mathrm{diag}(1, 1, -1)$ clearly has determinant -1. Another example, is the parity operation under which all coordinates change sign:

$$P : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} \tag{17}$$

One can show that any matrix satisfying $R^{\mathrm{T}} = R^{-1}$ with determinant -1 can be obtained by following the parity operation with a rotation. Thus, it suffices to consider parity alone.

A **scalar** is invariant under **both** rotations and parity while a **pseudo-scalar** is one that is invariant under rotations but changes sign under parity.

A **vector** is one that transforms identical to the position vector under both rotations as well as parity while a **pseudo-vector** transforms like the position vector under rotations but is invariant under parity.

The momentum $\vec{p}$ of a particle is a vector, while its angular momentum $\vec{L}$ is a pseudo-vector. This follows from the fact that both $\vec{r}$ and $\vec{p}$ are vectors and hence change sign under parity. Thus $\vec{L} = \vec{r} \times \vec{p}$ does *not* change sign, and hence is a pseudo-vector. In general, the cross-product of two vectors gives rise to a pseudo-vector, while the cross-product of a vector with a pseudo-vector is a vector. The dot product of two vectors is a scalar (e.g. the kinetic energy of a particle), while the dot product of a vector and a pseudo-vector is a pseudo-scalar. Vectors are sometimes called **polar vectors**, to distinguish them from **axial vectors** which is another name for pseudovectors.

The Lorentz force law

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

enables us to fix the behaviour of the electric and magnetic fields under parity. Both $\vec{F}$ and $\vec{v}$ are vectors and q is a scalar, it follows that **the electric field is a vector** while **the magnetic field is a pseudo-vector**. We usually decide if an object is a pseudo-vector or pseudo-scalar by considering simple known equations. **Never** equate a vector to a pseudo-vector or a scalar to a pseudo-scalar.

1.8 Form invariance of physical laws

The full power of working with objects which transform nicely under change of basis/rotations is that **the equations which describe the physical motion of particles can be written in manifestly invariant manner** i.e., they retain the same form in different bases. For instance, Newton's law

$$\vec{F} = m\vec{a}$$

is the same in all bases. The LHS is a vector and the RHS is the product of a scalar and a vector. Hence the RHS is a vector as well.

In more complicated situations, the equations need not be relations between vectors but those between tensors of the same rank. That way, invariance is guaranteed since both sides change in an identical fashion under change of bases and reflections. The simple way to verify this is to check that **free indices** are the same on both sides of any equation. One must also be cautious and check that the dummy (non-free) indices occur only twice. If any index occurs three or more timee, there's been a mistake somewhere.

1.9 Geometric meaning of the dot product and cross product

1.9.1 Area element as a vector

Given two vectors $\vec{u}$ and $\vec{v}$, one can construct a parallelogram with two edges $\vec{u}$ and $\vec{v}$(see Fig. 1.1). By considerations of elementary geometry, the area of the parallelogram is

$$\text{Area} = |\vec{u}||\vec{v}| \sin\theta \quad ,$$

where θ is the angle between the two vectors $\vec{u}$ and $\vec{v}$. Thus, one can see that the area can be written as Area $= |\vec{u} \times \vec{v}|$.

It turns out that it is better to think of the area of a parallelogram (a planar figure) as the vector $(\vec{u} \times \vec{v})$. Clearly, the direction of this vector is **normal** to the surface of the parallelogram. We thus define the **area vector** to be

$$\overrightarrow{\text{A}} = (\vec{u} \times \vec{v}) \quad . \tag{18}$$

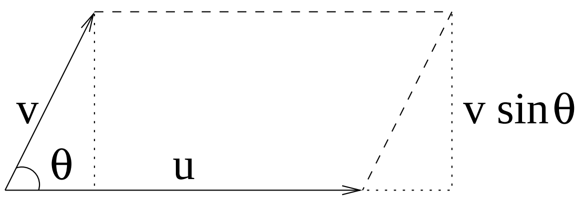

Figure 1: Area of a parallelogram

Note that there is an ambiguity of an overall sign in the definition, for we could also have chosen it to be $(\vec{v} \times \vec{u})$. We will return to this point when discussing the unit vector perpendicular to an area element.

1.9.2 Volume as a scalar

Just as two vectors can define a parallelogram, three (non co-planar) vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ can define a parallelepiped which has the three vectors as three of

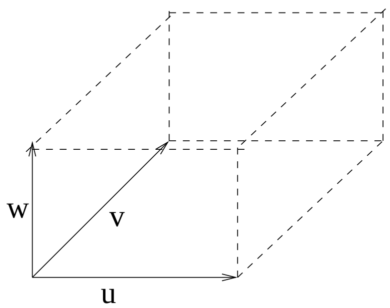

Figure 2: Volume of a parallelepiped

its edges (see Fig. 1.2). Again by means of elementary geometry, one can show that the volume of the parallelepiped is given by

$$\text{Volume} = \vec{u} \cdot (\vec{v} \times \vec{w}) \quad . \tag{19}$$

The above expression seems to treat the three vectors differently. but this is not so. We can re-express the above formula by using the Levi-Civita tensor

$$\text{Volume} = \epsilon_{ijk} u_i v_j w_k = \vec{u} \cdot (\vec{v} \times \vec{w}) = \vec{v} \cdot (\vec{w} \times \vec{u}) = \vec{w} \cdot (\vec{u} \times \vec{v}) \tag{20}$$

We note again that there is a sign ambiguity in the above definition. This can be removed by taking the magnitude in the RHS of the above equation, since we require the volume to be non-negative. Note that if $\vec{u}$, $\vec{v}$, $\vec{w}$ are co-planar, the volume of the parallelepiped vanishes automatically, as it must.
Exercise Obtain an explicit 3×3 matrix for an arbitrary rotation matrix $R(\hat{n}, \phi)$.
Recommended reading: V. Balakrishnan, *How is a vector rotated?*, Resonance, Vol. **4**, No. 10, pp. 61-68 (1999). Also available at the URL:

`http://www.physics.iitm.ac.in/%7Elabs/dynamical/pedagogy/`