An MHRD Project under its National Mission on Education throught ICT (NME-ICT)

A Gateway to All Post Graduate Courses

Subject: Physics

Production of Courseware

-Content for Post Graduate Courses

Paper No. : Electromagnetic Theory

Module : Mathematical Preliminaries - II

Development Team

Principal Investigator	Prof. Vinay Gupta, Department of Physics and Astrophysics, University of Delhi, Delhi
Paper Coordinator	Prof. V. K. Gupta, Department of Physics and Astrophysics,University of Delhi, Delhi
Content Writer	Prof. V. K. Gupta, Department of Physics and Astrophysics,University of Delhi, Delhi
Content Reviewer	Prof. Pankaj Sharan, Physics department from JMI, New Delhi

Description of Module	
Subject Name	Physics
Paper Name	Electromagnetic Theory
Module Name/Title	Mathematical Preliminaries - II
Module Id	M1

Contents

Learning Objectives:

From this module students may get to know about the following:

1. In this module we continue with study of vectors. Students come to know how, taking cue from position vector, vectors are defined in general on the basis of their properties under rotations.
2. The idea is then generalized to tensors of higher ranks based on the transformation properties under rotations
3. The concept of curvilinear coordinates as an alternative to Cartesian coordinates is introduced. The spherical polar and cylindrical coordinates are studied in detail.
4. The concept of Dirac delta function is introduced, both in one and three dimensions, and some of its important properties are described.

.

1. Vectors

1.1 Introduction to vectors

We will assume that you are familiar with the elementary definition of *vectors* as those physical quantities that have both magnitude and direction and add according to the parallelogram law of addition. Many physical quantities fall under this category, viz., displacement, velocity, momentum and force just to mention a few. We will denote a vector quantity by placing an arrow over it. For example, $\vec{x}, \vec{v}, \vec{P}, \vec{F}$ etc. The magnitude of a vector $\vec{A}$ is often written as $|\vec{A}|$, or sometimes simply as *A*. In contrast, *scalars* are physical quantities which have only a magnitude. Mass, energy, density and temperature are some of the examples.

Given two vectors $\vec{A}$ and $\vec{B}$ we define their *scalar product* or *dot product* as a scalar

$$\vec{A}.\vec{B} = |\vec{A}||\vec{B}|\cos(\theta)$$

where θ is the angle between the directions of the two vectors. If the two vectors are parallel, their scalar product is just the product of their magnitudes, $\vec{A}.\vec{B} = |\vec{A}||\vec{B}|$. If they are perpendicular, the scalar product vanishes, $\vec{A}.\vec{B} = 0$.

The *vector product* or *cross product* $\vec{A} \times \vec{B}$ of the two vectors is defined as

$$\vec{A} \times \vec{B} = |\vec{A}||\vec{B}|\sin(\theta)\hat{n}$$

Here θ is the *smaller* of the two angles between $\vec{A}$ and $\vec{B}$, $\hat{n}$ is a vector of unit magnitude which is normal to the plane containing $\vec{A}$ and $\vec{B}$, and its sense is given by the right hand screw rule. A unit vector is normally designated by putting a hat ($\hat{}$) over it. If the two vectors $\vec{A}$ and $\vec{B}$ are parallel, their cross product vanishes, $\vec{A}\times\vec{B}=0$. If they are perpendicular the magnitude of the cross product is just the product of their magnitudes, $|\vec{A}\times\vec{B}|=|\vec{A}||\vec{B}|$. In this case the three vectors, $\vec{A}$, $\vec{B}$ and $\vec{A}\times\vec{B}$ are *mutually perpendicular.*

- Note that whereas the scalar product of two vectors is *commutative*, i.e.,

$$\vec{A}.\vec{B}=\vec{B}.\vec{A},$$

the vector product is *anti-commutative*, i.e.,

$$\vec{A}\times\vec{B}=-\vec{B}\times\vec{A}$$

- From these definitions, one can prove many other relations involving the two kinds of products. For example

$$\begin{aligned}&\vec{A}.(\vec{B}\times\vec{C})=\vec{B}.(\vec{C}\times\vec{A})=\vec{C}.(\vec{A}\times\vec{B})\\&\vec{A}\times(\vec{B}\times\vec{C})=\vec{B}(\vec{C}.\vec{A})-\vec{C}(\vec{A}.\vec{B})\\&(\vec{A}\times\vec{B})\times(\vec{C}\times\vec{D})=(\vec{A}.\vec{C})(\vec{B}.\vec{D})-(\vec{A}.\vec{D})(\vec{B}.\vec{C})\end{aligned} \qquad (1)$$

1.1.1 Coordinate System

So far we have defined vectors without reference to any particular coordinate system. From a practical point of view it is often more useful to employ a definite coordinate system and work in terms of “components” of the vector. W set up a system of coordinates with an origin at some point *O* and three *mutually perpendicular* axes, *Ox, Oy, Oz.* Let $(\hat{x},\hat{y},\hat{z})$ be unit vectors along the three axes respectively. Then any vector $\vec{A}$ can be written as a sum of three vectors along the three axes:

$$\vec{A}=A_x\hat{x}+A_y\hat{y}+A_z\hat{z}\,.$$

A_x, A_y, A_z are the *components* of the vector $\vec{A}$ along the *x, y, z* directions, respectively. Geometrically they are the projections of the vector $\vec{A}$ on to *x, y, z* axes. Vectors add component-wise:

$$\vec{A}+\vec{B}=(A_x\hat{x}+A_y\hat{y}+A_z\hat{z})+(B_x\hat{x}+B_y\hat{y}+B_z\hat{z})$$
$$=(A_x+B_x)\hat{x}+(A_y+B_y)\hat{y}+(A_z+B_z)\hat{z}$$

Similarly, multiplication by a scalar is also component-wise:

$$k\vec{A}=k(A_x\hat{x}+A_y\hat{y}+A_z\hat{z})=kA_x\hat{x}+kA_y\hat{y}+kA_z\hat{z}$$

To write scalar and vector products using components, we first write down these products for the unit vectors:

$$\begin{aligned}
&\hat{x}.\hat{x}=\hat{y}.\hat{y}=\hat{z}.\hat{z}=1\\
&\hat{x}.\hat{y}=\hat{x}.\hat{z}=\hat{y}.\hat{z}=0\\
&\hat{x}\times\hat{x}=\hat{y}\times\hat{y}=\hat{z}\times\hat{z}=0\\
&\hat{x}\times\hat{y}=\hat{z},\qquad \hat{y}\times\hat{z}=\hat{x},\qquad \hat{z}\times\hat{x}=\hat{y}\\
&\hat{y}\times\hat{x}=-\hat{z},\qquad \hat{z}\times\hat{y}=-\hat{x},\qquad \hat{x}\times\hat{z}=-\hat{y}
\end{aligned}\qquad (2)$$

Using these elementary products, we can now write for any two vectors $\vec{A}$ and $\vec{B}$

$$\begin{aligned}
\vec{A}.\vec{B}&=(A_x\hat{x}+A_y\hat{y}+A_z\hat{z}).(B_x\hat{x}+B_y\hat{y}+B_z\hat{z})\\
&=A_xB_x+A_yB_y+A_zB_z
\end{aligned}\qquad (3)$$

and

$$\vec{A}\times\vec{B} = (A_x\hat{x} + A_y\hat{y} + A_z\hat{z})\times(B_x\hat{x} + B_y\hat{y} + B_z\hat{z})$$
$$= (A_yB_z - A_zB_y)\hat{x} + (A_zB_x - A_xB_z)\hat{y} + (A_xB_y - A_yB_x)\hat{z} \qquad (4)$$

This expression for the vector product can be written more neatly in a determinant form:

$$\vec{A}\times\vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \qquad (5)$$

Using these expressions for the two products, we can write the scalar triple product also in a determinant form:

$$\vec{A}.(\vec{B}\times\vec{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}. \qquad (6)$$

1.1.2 Position and Displacement Vectors

The position of a point in space can be described by fixing its coordinates *(x, y, z)*. The vector *to* that point *from* the origin is called the *position vector* of that point. For the position vector the notation $\vec{r}$ (and often $\vec{x}$) is employed:

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}. \qquad (7)$$

The distance from the origin is

$$r = \sqrt{x^2 + y^2 + z^2} = \sqrt{\vec{r}.\vec{r}}$$

The unit vector in the direction or the position vector $\vec{r}$ is

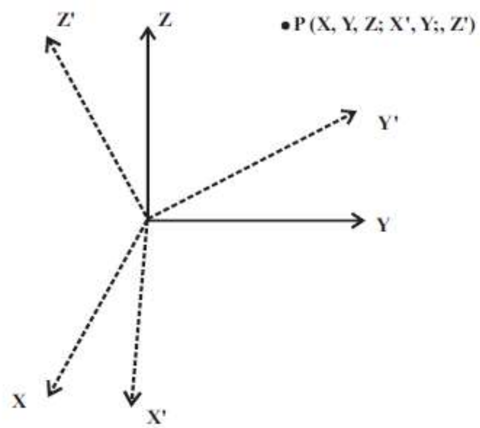
$$\hat{r} = \frac{\vec{r}}{r} = \frac{x\hat{x} + y\hat{y} + z\hat{z}}{\sqrt{x^2 + y^2 + z^2}} \qquad (8)$$

The *infinitesimal displacement vector* from (x, y, z) to $(x + dx, y + dy, z + dz)$, often written as $d\vec{r}$ (or $d\vec{x}$ or $d\vec{l}$) is

$$d\vec{r} = xd\hat{x} + dy\hat{y} + dz\hat{z}$$

1.2 Rotations and Vectors

Whereas the definition of a vector as a quantity having both magnitude and direction is adequate for many purposes, it is not entirely satisfactory. All quantities having magnitude and direction may not be vectors. A more satisfactory definition of vectors is in terms of their properties under certain transformations, viz., rotations. Consider a coordinate system (*Ox, Oy, Oz*). Let the coordinates of a point *P* in this system be (*x, y, z)*. What will be the coordinates of a point in the *rotated* coordinate system (*Ox′, Oy′, Oz′)*? The origin has remained the same, only the orientation has changed. Let the coordinates of a point *P*, whose position vector is $\vec{r}$, be (x, y, z) and (x', y', z') in the two systems. Then **[See Figure (to be drawn)]**



$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$$
$$= x'\hat{x}' + y'\hat{y}' + z'\hat{z}'$$

1.2.1 The tensor notation

The notation can be greatly simplified by representing the three axes as (x_1, x_2, x_3) instead of (x, y, z), and (x'_1, x'_2, x'_3) instead of (x', y', z'). Let the unit vectors of the unprimed system be represented by $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ and of the primed system by $(\hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$. The above relation can now be put in the form

$$\vec{r} = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3 = x'_1\hat{e}'_1 + x'_2\hat{e}'_2 + x'_3\hat{e}'_3 \qquad (9)$$

Let us take the dot product of both sides of the above equation with the unit vector $\hat{e}'_1$. Remember that like $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$, $(\hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$ are also mutually perpendicular. Hence

$$\vec{r}.e_1' = x_1(\hat{e}_1.\hat{e}'_1) + x_2(\hat{e}_2.e'_1) + x_3(\hat{e}_3.\hat{e}'_1) = x'_1(\hat{e}'_1.\hat{e}'_1) + x'_2(\hat{e}'_2.\hat{e}'_1) + x'_3(\hat{e}'_3.\hat{e}'_1)$$

or

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

(10)

If we denote the *direction cosines* $(\hat{e}_i'\hat{e}_j)$ between the x_i' and x_j axis by a_{ij}, then the above relation can be written as

$$x'_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$$

(11)

Similarly

$$x'_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$$
$$x'_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3$$

(12)

Together these equations can be written as

$$x'_i = \sum_{j=1}^{3} a_{ij}x_j$$

(13)

The transformation from the unprimed to the primed coordinate system is a linear homogeneous transformation.

There are nine coefficients, a_{ij} in all. However all nine are not independent. For transformation in a plane, say the *x*-*y* plane, out of the four coefficients, there is only one independent parameter, the angle of rotation. Since under a rotation the magnitude of a vector does not change, we have

$$r^2 = \vec{r}.\vec{r} = x_1^{\ 2} + x_2^{\ 2} + x_3^{\ 2} = x'_1{}^2 + x'_2{}^2 + x'_3{}^2$$

or

$$\sum_{i=1}^{3} x_i x_i = \sum_{i=1}^{3} x'_i\, x'_i \qquad (14)$$

Using (13) for x_i' in terms of x_i in (14) we have

$$\sum_{i=1}^{3} x'_i\, x'_i = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij} x_j \sum_{k=1}^{3} a_{ik} x_k$$

Hence

$$\sum_{i=1}^{3} x_i x_i = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij} x_j \sum_{k=1}^{3} a_{ik} x_k$$

The left hand side contains only the squared terms, $x_1^{\ 2}, x_2^{\ 2}, x_3^{\ 2}$, while the right hand contains cross terms as well. On comparing the coefficients of various terms we have

$$\sum_{i=1}^{3} a_{ij} a_{ik} = 1 \qquad if\ j = k$$

(15)

$$\sum_{i=1}^{3} a_{ij} a_{ik} = 0 \qquad if\ j \neq k$$

(16)

We can put the two results together by using a special symbol, called *Kronecker delta*, δ_{ij}. This quantity has value 1 if $i = j$ and value 0 if $i \neq j$. In terms of this symbol the two results can be combined as

$$\sum_{i=1}^{3} a_{ij} a_{ik} = \delta_{jk}$$

(17)

Since the expression is symmetric between *i* and *j*, this leads to six conditions on the nine coefficients, a_{ij}. Thus only three of them are independent.

The notation and writing can be further simplified by using what is called the *Einstein summation convention*. Under this convention

(i) An index (subscript) can appear **only once or twice**. Each index take values (*1, 2, 3*)

(ii) An index that appears once is a *free index*. Every free index represents a set of three equations for three values of the index.

(iii) An index that appears twice is a *dummy index.* The index is understood to be summed over, so that every such index represents a sum of three terms.

(iv) One dummy index can be freely replaced by another. Thus

$$\sum_{i=1}^{3} a_i a_i \equiv a_i a_i = a_j a_j = a_1^{\,2} + a_2^{\,2} + a_3^{\,2}$$

(v) The free indices must match on the two sides of an equation and on each term of an expression.

Using this notation, Equations (13), (14) and (17) can now be written as

$$x'_i = a_{ij} x_j \tag{18}$$

$$x_i x_i = x'_i\, x'_i \tag{19}$$

$$a_{ij} a_{ik} = \delta_{jk} \tag{20}$$

1.2.2 Matrix representation

Many of the equations above can be written in an even simpler form in terms of matrices. If we put the coordinates of a point in the form of a column matrix

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

and define the 3×3 square matrix of coefficients

$$a = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then equations(18) – (20) can be written as matrix equations

$$x' = ax \tag{21}$$

$$x^T x = x'^T x' \tag{22}$$

$$a^T a = I \tag{23}$$

The symbol *x* is being used for the set of three coordinates and not for the abstract vector which it represents. The superscript (*T*) refers to the transpose of a matrix and *I* refers to the unit matrix. The matrix *a* is such that

$$a^T = a^{-1}. \tag{24}$$

A matrix with this property is called an *orthogonal matrix*. Thus rotation is represented by an orthogonal matrix. This in fact is true not only in two or three dimensions but for rotations in any dimensions.

The inverse transformation from the primed to the unprimed coordinates is most easily obtained from inverting equation (21)

$$a^{-1}x' = a^{-1}ax = x$$

or

$$x = a^T x'$$
(25)

Thus the inverse transformation is given by the transpose of the coefficients.

1.3 Vector differential operators

In physics we deal not only with quantities which are scalars, vectors and tensors etc, but also with *scalar, vector and tensor fields*. The word "field" in this contest refers to any function of space and time coordinates. A charge produces an electric field at every point – it is a vector field $\vec{E}(x, y, z)$. So is also a potential associated with every point – it is a scalar field $V(x, y, z)$. We define certain *differential operators* that act on these fields and have definite transformation properties under rotation.

The *gradient operator* is a vector operator:

$$grad \equiv \vec{\nabla} \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$$

Acting on any scalar field $\Phi(\vec{x}) \equiv \Phi(x, y, z)$ it produces a vector field

$$grad\Phi \equiv \vec{\nabla}\Phi \equiv (\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z}).$$

The notation $\vec{\nabla}$ is more useful in tensor analysis. Thus

$$(\vec{\nabla})_i = (\frac{\partial}{\partial x_i}), i = 1,2,3$$

so that

$$(\vec{\nabla}\phi)_i = (\frac{\partial \phi}{\partial x_i}), i = 1,2,3$$

Similarly we define the *divergence* of a vector as

$$div\,\vec{E}(\vec{x}) \equiv \vec{\nabla}.\vec{E}(\vec{x}) \equiv (\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\frac{\partial E_z}{\partial z}) = \frac{\partial E_i}{\partial x_i}$$

and the $curl$ of a vector as

$$curl\,\vec{E}(\vec{x}) \equiv \vec{\nabla}\times\vec{E}(\vec{x}) \equiv (\frac{\partial}{\partial y}E_z - \frac{\partial}{\partial z}E_y)\hat{x} + (\frac{\partial}{\partial z}E_x - \frac{\partial}{\partial x}E_z)\hat{y} + (\frac{\partial}{\partial x}E_y - \frac{\partial}{\partial y}E_x)\hat{z}$$

In tensor notation it takes the form

$$(\vec{\nabla}\times\vec{E})_i = \varepsilon_{ijk}\frac{\partial}{\partial x_j}E_k$$

And then we have the *Laplacian operator*

$$\nabla^2 \equiv \vec{\nabla}.\vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial x_i \partial x_i}$$

The operator $\vec{\nabla}$ is a vector operator – it has all the properties of a differential operator as well as of a vector.

From these definitions one can easily prove the following identities

$$\vec{\nabla} \times \vec{\nabla} = 0$$

$$\vec{\nabla}.(\vec{\nabla} \times \vec{a}) = (\overline{\nabla} \times \vec{\nabla}).\vec{a} = 0$$

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{a}) = \overline{\nabla}(\vec{\nabla}.\vec{a}) - \nabla^2 \vec{a}$$

$$\vec{\nabla}.(\phi\vec{a}) = (\vec{a}.\vec{\nabla})\phi + \phi(\vec{\nabla}.\vec{a})$$

$$\vec{\nabla} \times (\phi\vec{a}) = \phi(\vec{\nabla} \times \vec{a}) - (\vec{a} \times \vec{\nabla})\phi$$

$$\vec{\nabla}(\vec{a}.\vec{b}) = (\vec{a}.\vec{\nabla})\vec{b} + (\vec{b}.\vec{\nabla})\vec{a} + \vec{a} \times (\vec{\nabla} \times \vec{b}) + \vec{b} \times (\vec{\nabla} \times \vec{a})$$

$$\vec{\nabla}.(\vec{a} \times \vec{b}) = \vec{b}.(\vec{\nabla} \times \vec{a}) - \vec{a}.(\vec{\nabla} \times \vec{b})$$

1.4 Integrals involving vectors

In electrodynamics, as in the rest of physics, we encounter many integrals, among which the most important are

***(i) Line Integral**:* $\int_a^b \vec{f}(\vec{x}).d\vec{l}$ over some specific path from *a* to *b*, or over a closed path $\oint \vec{f}.d\vec{l}$.

(ii) Surface Integral: $\int_S \vec{f}(\vec{x}).d\vec{a}$ over some specified open surface, or over a closed surface $\oint_S \vec{f}.d\vec{a}$

(iii)Volume Integral $\int_V \phi(\vec{x})d^3x$ over some volume.

1.4.1The Fundamental Theorems

There are certain theorems of fundamental importance that relate volume, surface and line integrals which are crucial to the subject of electrodynamics. We state below these theorems without proof:

(i) The Gauss theorem or divergence theorem: Given any vector function $\vec{A}$, the theorem states that

$$\int_V \vec{\nabla}.\vec{A}d^3x = \oint_S \vec{A}.\hat{n}d\vec{a}$$

Here *S* is the surface that encloses the volume *V* and $\hat{n}$ is the *outward drawn* unit normal to the surface. Equivalently, the relation can be written for a gradient of a scalar or for $curl(\vec{A})$:

$$\int_V \vec{\nabla}\phi d^3x = \oint_S \phi\hat{n}d\vec{a}$$

$$\int_V (\vec{\nabla}\times\vec{A})d^3x = \oint_S (\hat{n}\times\vec{A})d\vec{a}$$

(ii) The Green's first identity: It is a similar relation between volume and surface integrals:

$$\int_V (\phi\nabla^2\psi + \vec{\nabla}\phi.\vec{\nabla}\psi)d^3x = \oint_S \phi(\hat{n}.\vec{\nabla})\psi d\vec{a}$$

(iii) Stokes's theorem: The Stoke's theorem provides a relation between line integrals and surface integrals:

$$\oint_S (\vec{\nabla}\times\vec{A}).\hat{n}da = \oint_C \vec{A}.d\vec{l}$$

or equivalently

$$\oint_S (\hat{n}\times\vec{\nabla}\phi)da = \oint_C \phi d\vec{l}$$

Here *S* is an open surface and *C* is the contour around it and $d\vec{l}$ is a line element along it. The unit normal $\hat{n}$ to the surface *S* is defined by the right hand rule in relation to the direction of the line integral around it.

Summary

1. *In this module we have introduced the student to some of the essential mathematical preliminaries.*
2. *After giving the usual definition of vectors a more basic definition is provided which brings out the relation with rotations.*
3. *The tensor notation is introduced and its usefulness discussed.*
4. *The differential vector operators are introduced. Some of the fundamental theorems like Gauss theorem and Stoke's theorem which are crucial in electrodynamics are written out.*