UNIT 4 AN INTRODUCTION TO TENSORS

Structure

4.1 INTRODUCTION

In your study of elementary physics, you have learnt to classify physical quantities as scalars or vectors. However, in many physical problems, the notion of a vector becomes too restricted. For example, you have studied Ohm's law which states that 'the electric current in a medium is proportional to the applied electric field.' Mathematically, we can express it as: $\mathbf{j} = \sigma\mathbf{E}$ where **j** is the electric current density, **E** is the applied electric field and σ is a scalar representing electrical conductivity. Thus we can write $j_x = \sigma E_x$, $j_y = \sigma E_y$ and $j_z = \sigma E_z$ in a 3-D cartesian coordinate system.

But do you know that in this form the law applies only to isotropic media, i.e., media having the same properties in all *directions*? What do we do if a medium is not isotropic (i.e., anisotropic) as in crystals or a plasma in a magnetic field? For such media, **j** and **E** need not be in the same direction (Fig. 4.1). The current density in the x direction may depend on the electric field in all the three (x, y, z) directions. For low applied electric fields, we may have $j_1 = \sigma_{11}E_1 + \sigma_{12}E_2 + \sigma_{13}E_3$, $j_2 = \sigma_{21}E_1 + \sigma_{22}E_2 + \sigma_{23}E_3$, $j_3 = \sigma_{31}E_1 + \sigma_{32}E_2 + \sigma_{33}E_3$.

You can see that in such a situation, it is necessary to replace the scalar σ by a more general mathematical construct which when acting on **E** is capable of changing both its magnitude and direction. Such a construct is called a **tensor**. Generalising this idea, we say that if a relation $\mathbf{A} = B\,\mathbf{C}$ with **A** and **C** as non-parallel vectors holds in all orientations of a cartesian system, then B is a tensor.

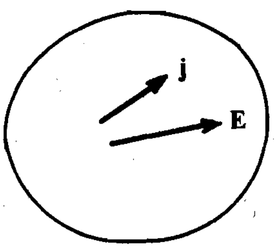

Fig.4.1: In an anisotropic medium, j and E are not parallel.

The coefficients $\sigma_{ij}\,(i, j = 1,2,3)$ are said to be the components of the **electrical conductivity tensor** of the medium. In a similar manner, we talk of **effective mass tensor** while dealing with motion of electrons and holes in semiconductors, **moment of inertia tensor** while dealing with the rotation of an irregular body, **susceptibility tensors** for electric and magnetic fields in an anisotropic medium etc. We shall discuss some of these examples in this unit. These and several other examples show that tensors are needed even in 'simple' classical physics. In relativistic physics, tensors are widely used in the special theory of relativity and much more so in the general theory. We shall mostly discuss tensors in classical Newtonian mechanics in this unit and restrict ourselves to observers who are stationary relative to each other.

Objectives

After studying this unit, you should be able to:

- define a tensor and explain the need of tensors in the formulation of scientific laws;
- explain the difference between contravariant and covariant components of vectors and of tensors; and
- determine whether a physical quantity transforms like a tensor.

4.2 WHAT IS A TENSOR?

In Newtonian mechanics, we say that physical quantities such as length, mass, energy, volume are *scalars* while velocity, force, linear momentum, electric field are *vectors*. Can you say why? To answer this question recall the concept of scalars and vectors that you have studied so far. In school you learnt that scalar quantities have magnitude only, whereas vector quantities have both magnitude and direction. While studying the physics elective PHE-04 you have refined these definitions. Now you know that the value of a scalar quantity does not change (or remains invariant) on changing the coordinate system. You have also learnt that a vector changes in a special way when it undergoes a coordinate transformation. Let us quickly revisit these concepts because *these form the basis for defining a tensor.*

4.2.1 Scalars and Vectors in Coordinate Transformations

Let us first consider a very simple case of two stationary observers S and S' situated in the two-dimensional cartesian coordinate systems (xy) and $(x'\,y')$ (Fig. 4.2). Both S and S' would measure the same value for the mass of a body or for the length of a rod etc. irrespective of their different coordinates. Conversely, we can say that *a physical quantity (represented by a single number) whose value is independent of coordinate transformations is a scalar.*

Thus if we denote by a and a' the values of some scalar quantity measured by two observers S and S', then we have

$$a' = a \qquad (4.1)$$

This is the crux of the matter. Merely the fact that something is represented only by a number does not qualify it to be called a scalar. In addition, this value should not depend on the coordinates chosen. This important property is known as *invariance* and we say that *a scalar is invariant under coordinate transformations.* All the physical quantities mentioned above such as mass, energy, power etc. are called *Galilean scalars* because they are invariant under Galilean transformations – transformations between two stationary coordinate systems in Newtonian mechanics.

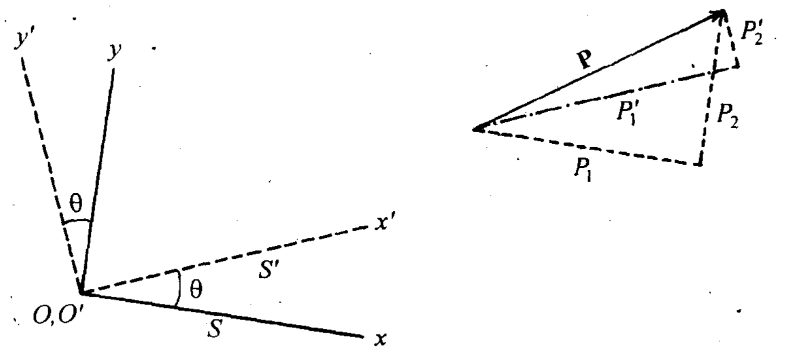

Fig.4.2: Two observers S and S' with their respective two-dimensional cartesian coordinate systems measuring the components of a vector P.

Now suppose S and S' want to measure a vector quantity (say velocity) in their respective coordinate system. Let us denote this vector by **P**. You know from Sec. 1.3.2, Unit 1 of PHE-4 that the components of **P** in $S(P_1, P_2)$ and $S'(P_1', P_2')$ are related by the equations:

$$P_1' = P_1 \cos\theta + P_2 \sin\theta \qquad (4.2a)$$

$$P_2' = -P_1 \sin\theta + P_2 \cos\theta \qquad (4.2b)$$

In terms of matrices, we can express Eqs. (4.2a and b) as

Lorentz Transformations

There is a purpose why we have been emphasizing 'Newtonian mechanics' and 'stationary observers' in the text. From Units 1 to 3 of PHE-11 you may know how and why special theory of relativity differs from Newtonian mechanics. This theory deals with a more general situation where observers may be in uniform relative motion with respect to each other. You know that the transformations connecting two spacetime coordinate systems (four coordinates x, y, z, t) which are in uniform relative motion with respect to each other are called *Lorentz transformations.* (Galilean transformations are a special case of Lorentz transformations when the relative velocity between the two observers is zero.) You know that the mass of an object and the length of a rod in the direction of relative motion between two observers do not remain invariant for such observers. Other such quantities which do not remain invariant are time interval, volume, energy, power etc. Thus many quantities which are Galilean scalars are **not** Lorentz scalars. The speed of electromagnetic radiation in vacuum, the rest mass of an object and the electrostatic charge are Lorentz scalars, which are found to have the same value for all observers in uniform relative motion with respect to each other.

$$\begin{pmatrix} P_1' \\ P_2' \end{pmatrix} = A \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.2c}$$

Squaring both sides of Eqs. (4.2a) and (4.2b) and adding, we get

$$P_1'^2 + P_2'^2 = P_1^2 + P_2^2 \tag{4.3}$$

The positive square root of each side of the above equation defines the *magnitude* of the 2-D vector **P**. Thus

$$P = |\mathbf{P}| = +\left(P_1^2 + P_2^2\right)^{1/2} \geq 0 \tag{4.4}$$

for observer S. Similarly we would have P' for observer S'. Then Eq. (4.3) reduces to

$$P' = P \tag{4.5}$$

All Numbers are not Scalars

Consider a 2-D or 3-D vector **P** with cartesian components (P_1, P_2) or (P_1, P_2, P_3) respectively, having magnitude P. We have seen above in Eq. (4.5) that P is a scalar. P_1, P_2 etc. are also numbers. Are they scalars? No, because these numbers do not remain invariant under coordinate transformations. We have seen in Eqs. (4.2a and b) how they transform from one cartesian system to another. Similarly, consider two observers S and S' in special theory of relativity moving with a uniform velocity with respect to each other. Let m and m' denote the mass of an object measured by S and S'. Both m and m' are numbers which are not invariant under Lorentz transformations. Therefore the physical quantity called the 'relativistic mass of an object' is not a Lorentz scalar but depends on the observer.

Let us extend this concept to a 3-D vector. You may now recall the definition of a 3-dimensional cartesian vector as given in Sec. 1.3.3 of Unit 1 of the course PHE-04:

A cartesian vector **P** is defined as a set of three numbers (components) in *every* rectangular coordinate system: If P_1, P_2, P_3 are the components in one system and P_1', P_2', P_3' are the components in a rotated system, these two sets of components are related by the equation:

$$\begin{pmatrix} P_1' \\ P_2' \\ P_3' \end{pmatrix} = A \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}, \tag{4.6a}$$

$$A = \left(a_{ij}\right) \tag{4.6b}$$

where a_{ij} are the cosines of the angles between the respective positive axes:

	x	y	z
x'	$a_{11} = \hat{\mathbf{i}}'.\hat{\mathbf{i}}$	$a_{12} = \hat{\mathbf{i}}'.\hat{\mathbf{j}}$	$a_{13} = \hat{\mathbf{i}}'.\hat{\mathbf{k}}$
y'	$a_{21} = \hat{\mathbf{j}}'.\hat{\mathbf{i}}$	$a_{22} = \hat{\mathbf{j}}'.\hat{\mathbf{j}}$	$a_{23} = \hat{\mathbf{j}}'.\hat{\mathbf{k}}$
z'	$a_{31} = \hat{\mathbf{k}}'.\hat{\mathbf{i}}$	$a_{32} = \hat{\mathbf{k}}'.\hat{\mathbf{j}}$	$a_{33} = \hat{\mathbf{k}}'.\hat{\mathbf{k}}$

(4.6c)

A physical quantity is called a **vector** if it transforms under a change of coordinate systems in accordance with Eqs. (4.6a) to (4.6c).

We would like to emphasize here that the components of any 3-D vector will transform from S to S' in the same manner as above. For example, the general vector **P** may stand for position vector, linear momentum, force, electric field, magnetic field, etc. In all these cases, its components P_i and P_i', as measured by two different observers, would be related to each other by the same equations, Eq. (4.6a) to Eq. (4.6c).

Let us now consider the coordinates of a point P being measured by S and S' of Fig. 4.2(a). Relative to S the coordinates of P are (x_1, x_2) while relative to S' these are (x_1', x_2') (see Fig. 4.3).

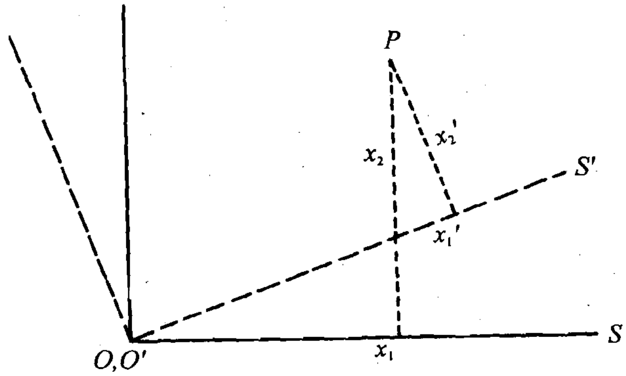

Fig.4.3: Observer S finds the coordinates of point P as (x_1, x_2) while S' finds them to be (x_1', x_2')

It is clear that these components are related to each other through the same equations, Eqs. (4.2a and b), with P_i replaced by x_i and P_i' by x_i'. We rewrite these for convenience:

$$x_1' = x_1 \cos\theta + x_2 \sin\theta,$$

$$x_2' = -x_1 \sin\theta + x_2 \cos\theta. \qquad (4.7a)$$

We can also write Eq. (4.7a) as

$$x_i' = \sum_{j=1}^{2} a_{ij} x_j, \quad i = 1,2. \qquad (4.7b)$$

There is another way of writing Eq. (4.7b) which simplifies the expression of tensors. To do so, we differentiate each primed coordinate x_i' with respect to the unprimed coordinates. Using Eqs. (4.7a) and (4.7b), we get

$$\frac{\partial x_1'}{\partial x_1} = \cos\theta = a_{11}, \quad \frac{\partial x_1'}{\partial x_2} = \sin\theta = a_{12},$$

$$\frac{\partial x_2'}{\partial x_1} = -\sin\theta = a_{21}, \quad \frac{\partial x_2'}{\partial x_2} = \cos\theta = a_{22}. \qquad (4.8a)$$

We can combine all the four equations in (4.8a) in the single equation

$$\frac{\partial x_i'}{\partial x_j} = a_{ij}, \qquad i, j = 1,2. \qquad (4.8b)$$

Similarly, for a 3-D coordinate transformation we have

$$\frac{\partial x_i'}{\partial x_j} = a_{ij}, \qquad i, j = 1,2,3 \qquad (4.9)$$

where a_{ij} are the elements of the transformation matrix defined in Eq. (4.6c).

Note that the coordinates of a system are independent of each other. For example, x_1 will be independent of x_2, x_3, …, etc. This is expressed mathematically as

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}, \qquad (4.10)$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \tag{4.11}$$

is the **Kronecker delta function.**

Now we use Eq. (4.8b) in Eq. (4.7b) and (4.2a). This gives

$$x_i' = \sum_{j=1}^{2} \frac{\partial x_i'}{\partial x_j} x_j \tag{4.12}$$

and

$$P_i' = \sum_{j=1}^{2} a_{ij} P_j = \sum_{j=1}^{2} \frac{\partial x_i'}{\partial x_j} P_j \tag{4.13}$$

We can generalize these results to a 3-D and N-D coordinate system so that j runs from 1 to 3 or 1 to N in Eqs. (4.12) and (4.13).

General Cartesian Transformations

Consider an N-dimensional Euclidean space and let x_i and x_i' ($i = 1,2, \ldots N$) be two coordinate systems in it. Since both coordinate systems are in the same space, each coordinate of one system is a function of the coordinates of the second system and vice versa:

$$x_i' = f(x_1, x_2, \ldots x_N) \tag{4.14}$$

and

$$x_i = g(x_1', x_2', \ldots x_N') \tag{4.15}$$

Now suppose that two observers S and S' in these cartesian coordinate systems x_i and $x_i'(1 \le i \le N)$ determine the components of a vector **P** as P_i and P_i', respectively. The transformation between them is given by Eq. (4.13) with coefficients a_{ij} given by Eq. (4.8b), except that now the indices i, j run over from 1 to N. Thus,

$$P_i' = \sum_{j=1}^{N} a_{ij} P_j = \sum_{j=1}^{N} \frac{\partial x_i'}{\partial x_j} P_j \tag{4.16}$$

Squaring both sides of the above equation, we get

$$P_i'^2 = \left(\sum_{j=1}^{N} a_{ij} P_j \right) \left(\sum_{k=1}^{N} a_{ik} P_k \right) = \sum_{j,k=1}^{N} a_{ij} a_{ik} P_j P_k \tag{4.17}$$

(Note that we *must* use two different summation indices on the right hand side. Do you know why? This is because we are multiplying two series.) Let us now sum both sides of Eq. (4.17) on i from 1 to N to get

$$\sum_{i=1}^{N} P_i'^2 = \sum_{j,k=1}^{N} \left(\sum_{i=1}^{N} a_{ij} a_{ik} \right) P_j P_k . \tag{4.18}$$

The norm P' of the vector $\mathbf{P}'$ is defined as

$$P'^2 = P_1'^2 + P_2'^2 + \ldots + P_N'^2 = \sum_{i=1}^{N} P_i'^2 \qquad (4.19)$$

A generalization of Eq. (4.5) to N-dimensional spaces requires that

$$P'^2 = P^2 = \sum_{i=1}^{N} P_i^2. \qquad (4.20)$$

A comparison of Eqs. (4.18), (4.19) and (4.20) tells us that

$$\sum_{j,k=1}^{N} \left(\sum_{i=1}^{N} a_{ij} a_{ik} \right) P_j P_k = \sum_{j=1}^{N} P_j^2. \qquad (4.21)$$

This is true if and only if

$$\sum_{i=1}^{N} a_{ij} a_{ik} = \delta_{jk}. \qquad (4.22)$$

Recall that a real, square matrix whose elements satisfy this property is called an *orthogonal* matrix. Thus we see that **the matrix A of transformation between two cartesian coordinate systems is an orthogonal matrix**. Eq. (4.22) tells us that the columns of matrix A are orthogonal to each other. It implies that its rows are orthogonal to each other. That is

$$\sum_{j=1}^{N} a_{ij} a_{kj} = \delta_{ik}. \qquad (4.23)$$

You may like to stop here and attempt an SAQ to understand these ideas better.

SAQ 1

Spend 5 min

Show that the dot product of two 3-D vectors is a scalar.

Now in your physics elective courses you have encountered a vector transformation that is slightly different from the one given by Eq. (4.16). The gradient of a scalar, $\nabla\phi$, is defined by

$$\nabla\phi = \hat{\mathbf{i}} \frac{\partial\phi}{\partial x_1} + \hat{\mathbf{j}} \frac{\partial\phi}{\partial x_2} + \hat{\mathbf{k}} \frac{\partial\phi}{\partial x_3} \qquad (4.24a)$$

Here we have used x_1, x_2, x_3 instead of x, y, z. How does $\nabla\phi$ transform under the coordinate transformation $(x_1, x_2, x_3) \rightarrow (x_1', x_2', x_3')$? The components of $\nabla\phi$ in the primed coordinate system are

$$\frac{\partial\phi'}{\partial x_i'} = \frac{\partial\phi}{\partial x_i'} \qquad \text{(Since } \phi \text{ is a scalar field, it is invariant.)}$$

Since x_i' are functions of x_i ($i = 1,2,3$), and vice versa we can use the chain rule for partial derivatives and write

$$\frac{\partial\phi'}{\partial x_i'} \equiv \frac{\partial\phi}{\partial x_i'} = \sum_{j=1}^{3} \frac{\partial\phi}{\partial x_j} \frac{\partial x_j}{\partial x_i'} \qquad (4.24b)$$

Notice that this differs from Eqs. (4.12) and (4.13) in that we have $\partial x_j / \partial x'_i$ instead of $\partial x'_i / \partial x_j$. This difference leads us to the definitions of **covariant** and **contravariant** vectors.

4.2.2 Covariant and Contravariant Vectors

By definition, **P** is called a **covariant vector** if its components transform as follows:

$$P'_i = \sum_{j=1}^{3} \frac{\partial x_j}{\partial x'_i} P_j \tag{4.25a}$$

Thus, the gradient of a scalar field is a **covariant vector**.

On the other hand, **P** is a **contravariant** vector if its components transform as follows:

$$P'_i = \sum_{j=1}^{3} \frac{\partial x'_i}{\partial x_j} P_j \tag{4.25b}$$

Notice that we can write the total differential of the primed coordinate as

$$dx'_i = \sum_j \frac{\partial x'_i}{\partial x_j} dx_j \tag{4.26}$$

Thus the components of a contravariant vector are simply the differentials of the coordinates.

The velocity and acceleration are examples of contravariant vectors as you can see in the following example.

Example 1: Velocity and acceleration as contravariant vectors

Differentiating x'_i with respect to t and making use of the chain rule of partial differentiation we get

$$\frac{dx'_i}{dt} = \sum_j \frac{\partial x'_i}{\partial x_j} \frac{dx_j}{dt} \tag{4.27}$$

The velocity components in S and S' are defined by

$$v'_i = \frac{dx'_i}{dt}, \tag{4.28a}$$

$$v_i = \frac{dx_i}{dt}, \tag{4.28b}$$

Hence Eq. (4.27) can be written as

$$v'_i = \sum_j \frac{\partial x'_i}{\partial x_j} v_j$$

On comparison with Eq. (4.25b) it is evident that the velocity **v** is a contravariant vector.

Differentiating Eq. (4.28a) with respect to time, we get

$$\frac{dv'_i}{dt} = \frac{d^2x'_i}{dt^2} = \sum_j \left(\frac{\partial x'_i}{\partial x_j}\right)\frac{d^2x_j}{dt^2} \text{ since} \left(\frac{\partial x'_i}{\partial x_j}\right) \text{is independent of time.}$$

Hence acceleration is also a contravariant vector.

It is customary to write the indices for contravariant vectors and tensors as superscripts rather than subscripts. For example, in this notation Eq. (4.16) for a contravariant vector becomes

$$P'^i = \sum_j \frac{\partial x'_i}{\partial x_j} P^j \tag{4.29}$$

You can see that for the transformation Eqs. (4.2 to 4.6) between *cartesian coordinate systems,*

$$\frac{\partial x'_i}{\partial x_j} = \frac{\partial x_j}{\partial x'_i} = a_{ij} \tag{4.30}$$

since both partial derivatives are equal to the cosine of the angle between the x'_i and x_j axis. However, this result does not hold for more general coordinates. For example, for the transformation from cartesian to spherical polar coordinates: $\frac{\partial x}{\partial \theta} \neq \frac{\partial \theta}{\partial x}$.

Thus, **in general there are two possible definitions of a vector (contravariant and covariant) which become identical for transformations between cartesian coordinate systems.**

At this stage you may be wondering as to what these vectors really mean? You may like to understand the physical or geometrical meaning of contravariant and covariant vectors. Strictly speaking, we should speak of **contravariant** and **covariant components,** rather than vectors. But this terminology is now used customarily. Let us explain these ideas with reference to the 2-D systems (x, y) and (r,θ). In the cartesian system, we can write

$$d\mathbf{r} = \hat{\mathbf{i}}\, dx + \hat{\mathbf{j}}\, dy$$

Consider a transformation from (x, y) to polar coordinates (r,θ):

$$d\mathbf{r} = \hat{\mathbf{i}}\, dx + \hat{\mathbf{j}}\, dy = \mathbf{e}_r ar + \mathbf{e}_\theta\, rd\theta \tag{4.31a}$$

where $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are the unit vectors in the polar coordinate system. We can also write $d\mathbf{r}$ as

$$d\mathbf{r} = \mathbf{a}_r dr + \mathbf{a}_\theta d\theta \tag{4.31b}$$

where $\mathbf{a}_r = \mathbf{e}_r$ and $\mathbf{a}_\theta = r\,\mathbf{e}_\theta$.

From Eq. (4.26) we know that dr, $d\theta$ (not $rd\theta$) are the **contravariant** components of $d\mathbf{r}$. Thus using the $\mathbf{a}_r$ and $\mathbf{a}_\theta$ vectors as basis vectors we have written $d\mathbf{r}$ in terms of its contravariant components. In a similar way we can write any vector **P** in terms of its *ordinary components* and unit **e** vectors or in terms of its *contravariant* components and the **a** vectors:

$$\mathbf{P} = P_r\, \mathbf{e}_r + P_\theta \mathbf{e}_\theta = P_r \mathbf{a}_r + \frac{P_\theta}{r}\mathbf{a}_\theta \tag{4.32}$$

Here P_r and $P_\theta\,(= rd\theta)$ are the ordinary components of **P**; P_r and $\frac{P_\theta}{r}$ are the contravariant components. Thus the contravariant components are the ordinary components *divided* by the scale factors (See Eqs. (3.36) to (3.44) of Unit 3, PHE-04). Now consider the components of ∇u in (x, y) and (r,θ) systems:

$$\mathbf{P} = \nabla u = \frac{\partial u}{\partial x}\hat{\mathbf{i}} + \frac{\partial u}{\partial y}\hat{\mathbf{j}} = \frac{\partial u}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial u}{\partial \theta}\mathbf{e}_\theta = P_r\mathbf{e}_r + P_\theta\mathbf{e}_\theta$$

We can also write

$$\mathbf{P} = \frac{\partial u}{\partial r}\mathbf{a}_r + \frac{\partial u}{\partial \theta}\mathbf{a}_\theta = P_r\mathbf{a}_r + r\,P_\theta\mathbf{a}_\theta$$

where $\mathbf{a}_r = \mathbf{e}_r$ and $\mathbf{a}_\theta = \mathbf{e}_\theta / r$.

Thus the *covariant* components of a vector are ordinary components *multiplied* by the scale factors. We can then write a vector in terms of its covariant components and the basis vectors $\mathbf{e}_r$ and $\mathbf{e}_\theta / r$. The **a** vectors are considered basic and the term covariant (vary with) is used to mean "vary in the same way as the basis **a** vectors". Then contravariant means to vary in the opposite way: the contravariant components are ordinary components *divided* by the scale factors. Note that the components of the vector and the basis vectors to be used with them always vary in the opposite ways so that the scale factors cancel.

You may wonder why we should be interested in these peculiar components and not stick to ordinary components. The reason lies in the simplicity of the transformation equations (4.25a) and (4.25b). When we say that a vector is covariant (or contravariant) what we really mean is that its covariant (contravariant) components are the simplest. For example, in polar coordinates, the contravariant components of $d\mathbf{r}$ are dr, $d\theta$ and its ordinary components are dr, $rd\theta$.

With these ideas in place, we will now define a tensor.

4.2.3 Defining a Tensor

Contravariant, covariant and mixed tensors of rank two are defined as follows:

$$A'^{ij} = \sum_{kl} \frac{\partial x'_i}{\partial x_k}\frac{\partial x'_j}{\partial x_l} A^{kl} \quad \text{contravariant tensor of rank two} \qquad (4.33a)$$

$$B'^{i}_{j} = \sum_{kl} \frac{\partial x'_i}{\partial x_k}\frac{\partial x_l}{\partial x'_j} B^{k}_{l} \quad \text{mixed tensor of rank two} \qquad (4.33b)$$

$$C'_{ij} = \sum_{kl} \frac{\partial x_k}{\partial x'_i}\frac{\partial x_l}{\partial x'_j} C_{kl} \quad \text{covariant tensor of rank two} \qquad (4.33c)$$

From the above definitions you can observe the following points:

- The rank of a tensor is determined by the number of partial derivatives (or direction cosines) in the definition: it is two for a second-rank tensor, three for a tensor of rank 3 and so on. Thus

  > A scalar is a tensor of rank zero. A (contravariant/covariant) vector is a (contravariant/covariant) tensor of rank one.

- The sum over each index (subscript or superscript) ranges over the dimensions of the space. Thus for a 2-D space k and l each take the values 1 and 2. For a 3-D space $k = 1, 2, 3$, $\ell = 1, 2, 3$ and so on.

- The number of indices (the rank of a tensor) is independent of the dimensions of the space.

- A^{kl} is contravariant with respect to both indices k and l; C_{kl} is covariant with respect to both indices and B_l^k is contravariant with respect to the index k but covariant with respect to the index l.

Remember that if we are using Cartesian coordinates, then all three forms of the tensors (contravariant, covariant and mixed) are the same.

Note that each of the above three equations contains two factors in the transformation coefficient. Each factor is a partial derivative of a coordinate of one system with respect to a coordinate of the other system.

In the case of a contravariant index, the transformation factor on the right hand side contains a *partial derivative of a new coordinate with respect to an old coordinate.*

In the case of a covariant index, the factor involves the *partial derivative of an old coördinate with respect to a new coordinate.*

The fact that contravariant indices are indicated as superscripts and covariant indices as subscripts is a matter of convention.

In a three-dimensional space, a contravariant second rank tensor A (with components A^{kl}) may be conveniently represented by writing out its components in a 3 × 3 array:

$$A = \begin{pmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{pmatrix} \qquad (4.34)$$

Do not confuse this representation with a matrix: Whereas a tensor can be represented as a matrix, any square array of numbers or functions does **not** form a tensor. The essential condition for a tensor is that its components transform according to Eq. (4.33a).

It is important for you to note the essential difference between a contravariant and a covariant tensor. In the case of a contravariant tensor, the tensor is represented by components along the *directions of coordinate increase*, whereas in the case of a covariant tensor, the tensor is represented by components in the *directions normal to constant-coordinate surfaces*. The physical examples discussed in this section will help in clarifying this point. The velocity and the acceleration vectors are represented in terms of components in the directions of coordinate increase, while the gradient vector is represented in terms of components in the directions orthogonal to the constant coordinate surfaces. In the case of a cartesian coordinate system, the coordinate direction x^i coincides with the direction orthogonal to the constant-x^i surface, so that the distinction between the covariant and the contravariant tensors vanishes.

You have studied about constant coordinate surfaces, unit vectors along the coordinate axes and normal to constant coordinate surfaces in Unit 3 of the course PHE-04. You might like to refresh your memory to understand this point.

It should be evident from Eqs. (4.31a) that the coordinate differentials dx_i are the components of a contravariant tensor of rank one – the infinitesimal displacement vector. It is important to note that the coordinates x^i, in spite of the notation, are *not* the components of a tensor.

Before we proceed further, let us introduce you to a summation convention followed in the study of tensors which simplifies the right hand side of Eqs. (4.33a to c).

Summation Convention

This convention was first used by Einstein: When an index (except N) is repeated in a term on one side of the equation, summation over it (from 1 to N) is implied. Thus we can write Eq. (4.33a) simply as

$$A'^{ij} = \frac{\partial x'_i}{\partial x_k}\frac{\partial x'_j}{\partial x_l} A^{kl}$$

Since k and l appear more than once on the right hand side, the sums over k and l run from 1 to N, where N defines the dimension of the space. This is the summation convention we will use from now on.

If an index appears only once in a term, it has a definite value – any value between 1 and N.

Such an index is called a **free index**. In Eq. (4.33a), i and j are free indices.

An index which is repeated and over which summation is implied is called a *dummy index*. In Eq. (4.33a), k and l are dummy indices. While using this summation convention, it is helpful to remember the following rules:

1. *A dummy index can be replaced by any other index which does not appear in the same term.*
2. *Two dummy indices appearing in a term can be interchanged with each other without any change in the expression.*

Let us consider an example to clarify these conventions.

Example 2

Let $a_i, b_i, c_i, d_i, 1 \le i \le N$, be four sets of N quantities each. Then, according to Einstein's summation convention, we can write

$$a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_N b_N$$

Similarly,

$$a_i b_j c_j = a_i b_1 c_1 + a_i b_2 c_2 + \ldots + a_i b_N c_N,$$

$$= a_i \left(b_1 c_1 + b_2 c_2 + \ldots + b_N c_N \right)$$

where i is a free index and has a fixed value between 1 and N. Moreover, it is obvious that

$$a_i b_i = a_j b_j = a_k b_k, \text{etc.}$$

Also,

$$a_i b_j c_j = a_i b_k c_k = a_i b_l c_l, \text{etc.}$$

You should note that in these equations, the dummy index j cannot be replaced by i because i already appears in the same term. Thus,

$$a_i b_j c_j \ne a_i b_i c_i. \qquad (4.35)$$

You can easily verify the validity of this inequality by noting that the left hand side of this equation is

$$a_i \left(b_1 c_1 + b_2 c_2 + \ldots + b_N c_N \right)$$

whereas its right hand side is

$$a_1 b_1 c_1 + a_2 b_2 c_2 + a_3 b_3 c_3 + \ldots$$

We will now illustrate the concept of a tensor with the help of an example. But you may like to try an SAQ first to get familiar with the summation convention.

SAQ 2

Spend 5 min

Write out in detail the following expressions, which are written here using summation convention:

(a) $a_{ij}\, x_i\, x_j$, in a 3-dimensional space

(b) $\dfrac{\partial u}{\partial x_i} \dfrac{\partial x_i}{\partial x'_j}$, in an N-dimensional space

---

Example 3: The electrical conductivity tensor

Let us verify whether the electrical conductivity tensor σ introduced in Sec. 4.1 transforms as a tensor of rank 2. To find the law of transformation for σ, let σ'_{kl} denote the components of σ in the primed coordinate system. Then

$$\mathbf{j}' = \sigma' \mathbf{E}'$$

Since $\mathbf{j}'$ and $\mathbf{E}'$ are vectors, they transform according to Eqs. (4.6a, b and c):

$$\mathbf{j}' = A\,\mathbf{j}, \qquad \mathbf{E}' = A\,\mathbf{E}$$

Hence, we have

$$A\,\mathbf{j} = \sigma' A\,\mathbf{E}$$

or

$$A\sigma\mathbf{E} = \sigma' A\,\mathbf{E} \qquad (\text{Since } \mathbf{j} = \sigma\mathbf{E})$$

Since this is true for all values of **E**, we have

$$A\sigma = \sigma' A$$

or

$$\sigma' = A\sigma A^{-1}$$

Since A is orthogonal, we can write

$$\sigma' = A\sigma A^{T}$$

In component form, we can write this equation as

$$\sigma'_{kl} = \sum_i \sum_j a_{ki}\sigma_{ij}a_{lj} = a_{ki}a_{lj}\sigma_{ij} \quad \text{using the summation convention}$$

Thus, as per Eq. (4.33a), σ transforms as a contravariant tensor of rank two.

You have already seen that velocity and acceleration are tensors of rank one. Let us now consider tensors of higher ranks.

4.2.4 Tensors of Higher Ranks

In general, a tensor of rank r in an N-dimensional space is a set of N components which transform from one coordinate system to another according to specific laws, known as *tensor transformation laws*. These transformation laws are similar to Eqs. (4.33a) to (4.33c) except that they will involve the product of r partial derivatives of one set of coordinates with respect to another set instead of just two. Thus the components of a contravariant tensor of rank 3 would transform according to

$$A'^{\alpha\beta\gamma} = \frac{\partial x'_\alpha}{\partial x_i}\frac{\partial x'_\beta}{\partial x_j}\frac{\partial x'_\gamma}{\partial x_k} A^{ijk}, \qquad \text{contravariant tensor of rank 3} \qquad (4.36a)$$

whereas components of a mixed tensor of rank 3 transform according to

$$A'^{\alpha\beta}_{\gamma} = \frac{\partial x'_\alpha}{\partial x_i}\frac{\partial x'_\beta}{\partial x_j}\frac{\partial x_k}{\partial x'_\gamma} A^{ij}_{k}, \qquad (4.36b)$$

mixed tensors of rank 3

$$A'^{\alpha}_{\beta\gamma} = \frac{\partial x'_\alpha}{\partial x_i}\frac{\partial x_j}{\partial x'_\beta}\frac{\partial x_k}{\partial x'_\gamma} A^{i}_{jk}, \qquad (4.36c)$$

and for a covariant tensor, we have

$$A'_{\alpha\beta\gamma} = \frac{\partial x_i}{\partial x'_\alpha}\frac{\partial x_j}{\partial x'_\beta}\frac{\partial x_k}{\partial x'_\gamma} A_{ijk}, \qquad \text{covariant tensor of rank 3} \qquad (4.36d)$$

So you see that for a tensor of rank 3, we could have different types of components: the number of contravariant and covariant indices could be 3+0, 2+1, 1+2 or 0+3; the total number of indices in each case being 3. Each of these tensors would have N^3 components, N being the dimensionality of space.

The dimensionality of space is $N = 3$ in Newtonian mechanics and $N = 4$ in the theories of relativity. These two values occur most frequently in physics although other values also occur in certain theories. In Table 4.1 we have listed the number of components of tensors of various ranks in these two cases.

Table 4.1: Number of components of tensors of different ranks in three and four-dimensional spaces

Dimension N	Rank of tensor r					
	0	1	2	3	4	5
3	1	3	9	27	81	243
4	1	4	16	64	256	1024

From Table 4.1 we will note that a tensor of rank 0 has only one component in a space of any dimensionality. According to the general transformation law, there would be no factor involving the partial derivatives. Therefore the only component of this tensor would retain the same value or the same functional form in any coordinate system. That is, a tensor of rank zero is invariant under coordinate transformations and is therefore a *scalar*.

In classical physics, mass, length and volume, energy, electronic charge, speed of electromagnetic radiation in vacuum etc. are scalar quantities. These physical quantities do not depend on the choice of the coordinate system. In the 4-D special theory of relativity, mass, length, volume, energy etc. depend on the observer and the coordinate system. But electronic charge, speed of electromagnetic radiation in vacuum etc. do not depend on the observer even in 4-D space-time. Hence they are scalars in special theory of relativity too.

In the case of a tensor of rank 1, it has as many components as the dimensionality of space. For example, in Newtonian mechanics, force, velocity, linear momentum, acceleration, angular momentum etc. are represented by tensors of rank 1, that is *vectors*. In 4-D space-time, we have the concept of four vectors, such as energy-momentum four-vector, velocity four-vector, charge density-current density four vector etc. Each of these has four components which transform according to tensor of rank 1.

It is good to always remember that *tensors of rank zero are scalars while tensors of rank* 1 *are vectors.*

SAQ 3

Spend 10 min

Show that the Kronecker delta function is a mixed tensor of rank 2.

You know that physics deals with cause-effect relationships. For example, an electric field (cause) gives rise to electric current density (effect), stress (cause) leads to strain (effect), angular velocity (cause) results in angular momentum (effect), etc. Each of these relationships has to be expressed in a tensor form. As a matter of fact, all quantities appearing in the mathematical formulation of physical laws must be tensors.

The three quantities that appear in a physical relationship are *cause*, *effect*, and the *coefficient* connecting them. There is a general rule which says:

If cause is a tensor of rank m and effect is a tensor of rank n, then the coefficient connecting them must be a tensor of rank m+n.

In Example 3, the current density **j** and the electric field **E** are tensors of rank 1 (vectors). So the electric conductivity is a tensor of rank 2. We shall consider some more examples here. We shall use the summation convention and also the fact that the distinction between contravariant and covariant components vanishes in the case of cartesian coordinates, allowing us to express all indices as covariant indices.

Example 4: Higher rank tensors in physics

Moment of inertia tensor

Both angular velocity ω and angular momentum **L** of a body are tensors of rank 1 (vectors). Hence the coefficient connecting them must be a second rank tensor. Thus the simple elementary relationship $\mathbf{L} = I\omega$, which is valid only when **L** and ω are parallel to each other, must be modified to

$$L_p = I_{pq}\omega_q \tag{4.37}$$

where I_{pq} is the *moment of inertia tensor* of rank 2. It can be expressed as a (3×3) matrix about which you have studied in Unit 3.

Susceptibility tensors

For anisotropic media, the electric polarization **P** and electric field vector **E** are tensors of rank 1 related by

$$P_k = \varepsilon_{kl} E_l \tag{4.38}$$

Similarly, the intensity of magnetization **I** and magnetic field strength **H** are tensors of rank 1:

$$I_k = \mu_{kl} H_l \tag{4.39}$$

The electric susceptibility ε_{kl} and magnetic susceptibility μ_{kl} are tensors of rank 2.

Tensor of elastic constants

Both the elastic stress and elastic strain are really tensors of rank 2. Therefore, when a body is subjected to elastic stress described by the tensor X_{ij}, the elastic strain developed in the body described by the tensor e_{ij} is given by

$$e_{ij} = S_{ijkl} X_{kl}, \tag{4.40}$$

where S_{ijkl} is a tensor of rank 4 called the *tensor of elastic constants*. The familiar Young's modulus, Poisson ratio, modulus of rigidity etc. are typical examples of these elastic constants.

Piezoelectric strain tensor

Some crystals develop an electric polarization when they are subjected to mechanical stress Such crystals are known as *piezoelectric crystals*. Since stress X_{ij} is a second rank tensor and electric polarization P_i is a first rank tensor (vector), the coefficient connecting them is a third rank tensor known as the *piezoelectric strain tensor*:

$$P_i = d_{ijk} X_{jk} \tag{4.41}$$

You must be familiar with domestic gas lighters which use a quartz crystal, quartz being a piezoelectric material. When we press the knob, a stress is developed on the crystal, which produces an electric field across its faces, which in turn gives rise to a spark across the metallic gap. (You must not confuse this with the electric gas lighter which works on electrical mains supply and which does not need a piezoelectric crystal.) In the reverse action, if an alternating current is fed to a piezoelectric crystal, it elongates and contracts alternately and can be used as an ultrasonic generator.

Curvature tensor

The curvature tensor which describes the spacetime geometry in the general theory of relativity is a tensor of rank 4 (in 4-D spacetime) denoted by R_{ijkl}. We shall not go into more details here.

The order in which the indices appear in our description of a tensor is important. In general, A^{mn} is independent of A^{nm}. However, there are two cases of special interest, namely, symmetric and antisymmetric tensors which we will discuss now.

4.2.5 Symmetric and Antisymmetric Tensors

If

$$A^{ji} = A^{ij} \quad \text{or} \quad B_{ji} = B_{ij}, \quad 1 \le i,\ j \le N. \tag{4.42a}$$

a tensor is said to be **symmetric**.

On the other hand, the tensor is said to be **antisymmetric** if

$$A^{ji} = -A^{ij} \quad \text{or} \quad B_{ji} = -B_{ij}, \quad 1 \le i,\ j \le N. \tag{4.42b}$$

The symmetry and antisymmetry property of tensors is similar to those of matrices. (See Unit 2 of this course.) The property of symmetry (or antisymmetry) of a tensor between two contravariant or two covariant indices is invariant under coordinate transformations. That is, if the tensor is symmetric (or antisymmetric) in one coordinate system, it will be so in any other coordinate system. Let us prove this statement.

Let A^{ij} be a symmetric tensor in one coordinate system, so that

$$A^{ji} = A^{ij}.$$

Now we wish to examine whether its components $A'^{\alpha\beta}$ in some other coordinate system will also show this property.

Let us repeat Eq. (4.33a) for transformation of A^{ij} for ready reference:

$$A'^{\alpha\beta} = \frac{\partial x'_\alpha}{\partial x_i} \frac{\partial x'_\beta}{\partial x_j} A^{ij}.$$

We want to check whether $A'^{\beta\alpha}$ is equal to $A'^{\alpha\beta}$ or not. The component $A'^{\beta\alpha}$ is given by

$$A'^{\beta\alpha} = \frac{\partial x'_\beta}{\partial x_i} \frac{\partial x'_\alpha}{\partial x_j} A^{ij}. \tag{4.43}$$

If we interchange the dummy indices i and j in the above equation (see Rule 2 before Example 2) and use Eq. (4.42a), we have

$$A'^{\beta\alpha} = \frac{\partial x'_\beta}{\partial x_j}\frac{\partial x'_\alpha}{\partial x_i} A^{ji}$$
$$= \frac{\partial x'_\beta}{\partial x_j}\frac{\partial x'_\alpha}{\partial x_i} A^{ij} = A'^{\alpha\beta}, \quad (4.44)$$

in accordance with Eq. (4.42a). This proves the statement.

However, the 'symmetry' of a tensor between one covariant and one contravariant index is not an invariant property. (See Terminal Question 2.)

A tensor of rank higher than 2 may also show symmetry property between two similar indices. The stress tensor X_{kl} and the strain tensor e_{ij} mentioned in Eq. (4.40) are symmetric tensors. Therefore the tensor S_{ijkl} of elastic constants has the symmetry property

$$S_{ijkl} = S_{jikl} = S_{ijlk} = S_{jilk}. \quad (4.45)$$

For the same reason, the piezoelectric strain coefficient tensor d_{ijk} of Eq. (4.41) is symmetric in its second and third indices, that is

$$d_{ijk} = d_{ikj}. \quad (4.46)$$

The curvature tensor R_{ijkl} mentioned in Example 4 is antisymmetric between the first two indices as well as its third and fourth indices. Thus,

$$R_{jikl} = -R_{ijkl}, R_{ijlk} = -R_{ijkl}.$$
$$\Rightarrow R_{jilk} = R_{ijkl}. \quad (4.47)$$

SAQ 4

Spend 10 min

T_{ikl} is antisymmetric with respect to all pairs of indices. Express this statement mathematically.

To avoid errors arising due to wrong usage of covariant and contravariant indices as well as free and dummy indices, it is necessary to make sure that every tensor equation and every term involving a tensor satisfies certain criteria. There are some simple rules for checking the correctness of indices in a tensor equation. We have explained these below.

1. *A free index should match in all terms throughout the equation.* This means that if a free index occurs as a contravariant (covariant) index in one term, it should occur as a contravariant (covariant) index in each term of the equation.

2. *A dummy index should match in each term of the equation separately.* A dummy index may occur only in some of the terms of an equation. When it occurs in a term, it should occur twice, once in a contravariant position and once in a covariant position.

3. *No index should occur more than twice in any term.*

4. When a coordinate differential such as ∂x_i occurs in a term, i is to be regarded as a contravariant index if ∂x_i occurs in the numerator and as a covariant index if it occurs in the denominator. Thus, in an expression such as $\partial x_i / \partial x_\alpha$, i is a contravariant index while α is a covariant index.

You are aware of the principle of relativity. Physical laws are invariant with respect to an observer – they are independent of the position of the observer (translation of the coordinate system) and of the orientation of the observer's frame of reference (rotation of the coordinate system).

Owing to this, in the mathematical formulation of physical laws, only those quantities which possess such invariance properties can occur. These are scalars, vectors, and in general, tensors. Any other quantity not possessing these invariance properties just cannot occur in the mathematical representation of physical laws.

We must emphasize that a tensor is an entity which is independent of the system of coordinates to which its components are referred. If a quantity expressed as a tensor exists in one coordinate system, it exists in all coordinate systems and is not just a consequence of the *choice* of a coordinate system as are centrifugal force and coriolis force in Newtonian mechanics. It is only its components which undergo transformation from one coordinate system to another. The tensor notation using subscripts and superscripts has the disadvantage of not emphasizing this very important aspect. On the other hand, it has the advantage that the rank of the tensor and its free and dummy indices are immediately revealed. This is clear with reference to vectors which are tensors of rank unity and in which case two notations are commonly used. Thus we may denote a vector by **u** or u_i. While the components u_i depend on the coordinate system chosen, the vector as such (that is **u**) is an invariant quantity independent of the coordinate system.

It is also interesting to compare the way scalars, vectors and tensors are introduced in elementary physics textbooks. Scalars are generally introduced by saying that they are physical quantities which can be specified by one number only, and vectors by saying that they are physical quantities possessing both magnitude and direction. Then one goes on to define addition, subtraction, scalar product of vectors, etc. Rigorously, a physical scalar is a quantity which remains invariant under all coordinate transformations. Similarly, the rigorous way to define a vector is to define a vector space first and then to say that elements of a vector space are called vectors as is done in the Mathematics elective MTE-02. One could then introduce transformations of vectors in a vector space. The definition of a vector space includes within it vector addition. Thus the parallelogram law of vectors, which is familiar form of vector addition in three dimensions, is *not a property* of vectors but their *defining characteristic*. The position of a scalar as a point on a straightline or in a complex plane, and the direction and magnitude of a vector are only their geometrical or graphical representations, which help us to understand some of the operations and why we are carrying them out.

In the case of tensors, it is not possible (or at least not easy) to make any geometrical pictures, and *hence tensors have to be introduced only through their transformations under changes of the coordinate system*. Presented in this fashion you may have found it a little difficult to grasp the concept of tensors. However, this presentation has its own advantage and merit, as it saves us from any confusion and stops any wrong ideas from entering into our heads. Moreover, even in quantum mechanics, physical pictures of electronic orbits in an atom or of the spin angular momentum have become meaningless, and we have to live with abstract operators, eigenvalue equations and transition probabilities.

Let us now summarise what you have studied in this unit.

4.3 SUMMARY

- **Every physical quantity is really a tensor**. For example, length, mass, force, velocity, electrical conductivity, elastic strain and stress etc. Every law in physics is a relation among physical quantities represented by tensors.

- The **components** of a tensor transform from one coordinate system to another in specific ways. For example, a tensor of rank 2 obeys the following transformation laws:

$$A'^{ij} = \sum_{kl} \frac{\partial x'_i}{\partial x_k} \frac{\partial x'_j}{\partial x_l} A^{kl} \qquad \text{contravariant tensor of rank two}$$

$$B'^{i}_{j} = \sum_{kl} \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} B^k_l \qquad \text{mixed tensor of rank two}$$

$$C'_{ij} = \sum_{kl} \frac{\partial x_k}{\partial x'_i} \frac{\partial x_l}{\partial x'_j} C_{kl} \qquad \text{covariant tensor of rank two}$$

- A tensor of **rank** r in an N-dimensional space has N^r components. A tensor of rank r has r free indices, each of which can take values from 1 to N.

- Scalars and vectors are special cases of a tensor. A **tensor of rank** 0 is a **scalar** while a **tensor of rank** 1 is a **vector**.

- Tensor indices are of two types, **contravariant** and **covariant**. The *distinction between these two vanishes in a cartesian coordinate system*, but is important in any other coordinate system.

4.4 TERMINAL QUESTIONS

Spend 30 *min*

1. The Riemann curvature tensor used by Einstein to describe the non-Euclidean geometry of spacetime is a tensor of rank 4. Write down the transformation of its components from one coordinate system to another when (a) the tensor has contravariant rank 3 and covariant rank 1, and (b) the tensor has covariant rank 4.

2. The components of a mixed tensor A^i_j show symmetry in one coordinate system, that is $A^j_i = A^i_j$. Examine whether the tensor will retain this property in any other coordinate system.

3. If A^{ij} is an antisymmetric tensor and B_i is a vector, show that $A^{ij} B_i B_j = 0$ (summation convention assumed).

4.5 SOLUTIONS AND ANSWERS

Self-assessment questions

1. We can express the dot product of two 3-D vectors **a** and **b** as

$$\sum_{i=1}^{3} a_i b_i$$

In matrix form we can write this as

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

The components a_i and b_i transform as

$$\begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \end{pmatrix} = A \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} b'_1 \\ b'_2 \\ b'_3 \end{pmatrix} = A \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Thus in the primed coordinate system the product $\sum_i a_i' \, b_i'$ becomes

$$\begin{pmatrix} a_1' & a_2' & a_3' \end{pmatrix} \begin{pmatrix} b_1' \\ b_2' \\ b_3' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} A^T A \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Since A is orthogonal $A^T A = I$ and hence

$$\sum_i a_i' \; b_i' = \sum_i a_i \; b_i$$

Thus, the scalar product is invariant under coordinate transformations and is a scalar.

2. a) $$a_{ij} \, x_i \, x_j = \sum_{i,j=1}^{3} a_{ij} \, x_i \, x_j,$$

since both i and j are dummy indices and $i,j = 1,2,3$ for a 3-D space. Thus

$$\begin{aligned} a_{ij} x_i x_j = a_{11} x_1^2 + a_{12} x_1 x_2 + a_{13} x_1 x_3 \\ + a_{21} x_2 x_1 + a_{22} x_2^2 + a_{23} x_2 x_3 \\ + a_{31} x_3 x_1 + a_{32} x_2 + a_{33} x_3^2 \end{aligned}$$

b) $$\frac{\partial u}{\partial x_i} \frac{\partial x_i}{\partial x_j'} = \frac{\partial u}{\partial x_1} \frac{\partial x_1}{\partial x_j'} + \frac{\partial u}{\partial x_2} \frac{\partial x_2}{\partial x_j'} + \ldots + \frac{\partial u}{\partial x_N} \frac{\partial x_N}{\partial x_j'}$$

3. The question is does δ_i^j transform according to Eq. (4.33b)? By the definition of Kronecker delta and using the summation convention we can write

$$\delta_i^j \frac{\partial x_k'}{\partial x_j} \frac{\partial x_i}{\partial x_l'} = \frac{\partial x_k'}{\partial x_j} \frac{\partial x_j}{\partial x_l'} \qquad \text{since } \delta_i^j = 1 \text{ only for } i = j$$

Now

$$\frac{\partial x_k'}{\partial x_j} \frac{\partial x_j}{\partial x_l'} = \frac{\partial x_k'}{\partial x_l'} \qquad \text{by chain rule for the right hand side.}$$

But x_k' and x_l' are independent coordinates. Hence they will be non-zero only if $k = l$ i.e., $\frac{\partial x_k'}{\partial x_l'} = \delta_l'^k$. Hence

$$\delta_l'^k = \frac{\partial x_k'}{\partial x_j} \frac{\partial x_i}{\partial x_l'} \delta_i^j$$

implying that δ_i^j is a mixed tensor of rank 2.

4. $$T_{ikl} = -T_{kil} = T_{kli} = -T_{lki}$$

$$T_{ikl} = -T_{ilk} = T_{lik} = -T_{lki}$$

Terminal Questions

1. If the components in the first coordinate system are denoted by unprimed symbols and those in the new coordinate system are denoted by primed symbols, the transformation in the two cases would be given by

(a) $$R'^{\alpha\beta\gamma}_{\sigma} = \frac{\partial x'^{\alpha}}{\partial x^{i}} \frac{\partial x'^{\beta}}{\partial x^{j}} \frac{\partial x'^{\gamma}}{\partial x^{k}} \frac{\partial x^{l}}{\partial x'^{\sigma}} R^{ijk}_{l};$$

(b) $$R'_{\alpha\beta\gamma\sigma} = \frac{\partial x^{i}}{\partial x'^{\alpha}} \frac{\partial x^{j}}{\partial x'^{\beta}} \frac{\partial x^{k}}{\partial x'^{\gamma}} \frac{\partial x^{l}}{\partial x'^{\sigma}} R_{ijkl}.$$

2. The transformation of the tensor components is

$$A'^{\alpha}_{\beta} = \frac{\partial x'^{\alpha}}{\partial x^{i}} \frac{\partial x^{j}}{\partial x'^{\beta}} A^{i}_{j}.$$

so

$$A'^{\beta}_{\alpha} = \frac{\partial x'^{\beta}}{\partial x^{i}} \frac{\partial x^{j}}{\partial x'^{\alpha}} A^{i}_{j}$$

$$= \frac{\partial x'^{\beta}}{\partial x^{i}} \frac{\partial x^{j}}{\partial x'^{\alpha}} A^{j}_{i} \qquad \text{(because } A^{j}_{i} = A^{i}_{j}\text{)}$$

$$= \frac{\partial x'^{\beta}}{\partial x^{j}} \frac{\partial x^{i}}{\partial x'^{\alpha}} A^{j}_{i} \qquad \text{(interchange of dummy indices)}$$

$$\neq A'^{\alpha}_{\beta}.$$

Thus the 'symmetry' of a tensor between one covariant and one contravariant index may be a freak, incidental property in one coordinate system. It is not a universal, invariant property in all coordinate systems.

3. Starting with the left hand side and using various rules described in this unit, we have

$$A^{ij} B_i B_j = A^{ji} B_j B_i \qquad \text{(interchange of dummy indices)}$$

$$= -A^{ij} B_j B_i \qquad \text{(antisymmetry of } A^{ij}\text{)}$$

$$= -A^{ij} B_i B_j,$$

where we have used the fact that B_i and B_j commute with each other since they are simple components, not operators. The above equation shows that the expressions on the left and the right hand sides are the same except for a change of sign. Transferring the right hand side to the left, we have

$$2A^{ij} B_i B_j = 0 \Rightarrow A^{ij} B_i B_j = 0.$$