4

Vector Transformation and Cartesian Tensors

The universal validity of physical laws is best expressed in terms of mathematical quantities that are independent of any reference frame. Yet physical problems governed by these laws can be solved, in most cases, only if the relevant quantities are resolved into their components in some coordinate system. For example, if we consider a block sliding on an inclined plane, the motion of the block is of course governed by Newton's second law of dynamics $\mathbf{F} = m\mathbf{a}$, in which no coordinates appear. However, to get the actual values of velocity, acceleration, etc. of the block, we have to set a coordinate system. We will obtain the correct answer no matter how we orient the axes, although some orientations are more convenient than others. It is possible to take the x-axis horizontal or along the incline. The components of $\mathbf{F}$ and $\mathbf{a}$ in the x and y directions are of course different in these two cases, but they will combine to give the same correct results. In other words, if the coordinate system is rotated, the components of a vector will of course change. But they must change in a specific way in order for the vector equation to remain valid. For this reason, the vector field is best defined in terms of the behavior of its components under axes rotation.

When the coordinate system is rotated, the transformation of the components of the position vector $\mathbf{r}$ can be expressed in terms of a rotation matrix. We will use this rotation matrix to define all other vectors. The properties of this rotation matrix will be used to analyze a variety of ways of combining the components of two or more vectors. This approach to vector analysis can be easily generalized to vectors of dimensions higher than three. It also naturally leads to tensor analysis.

Many physical quantities are neither vectors nor scalars. For example, the electric current density $\mathbf{J}$ flowing in a material is linearly related to the electric field $\mathbf{E}$ that drives it. If the material is isotropic, the three components of $\mathbf{J}$ and $\mathbf{E}$ are related by the same constant σ in the Ohm's law

$$J_i = \sigma E_i,$$

where σ is known as the conductivity. However, if the material is anisotropic (nonisotropic), the direction of the current is different from the direction of the field. In that case, the Ohm's law is described by

$$J_i = \sum_{j=1}^{3} \sigma_{ij} E_j,$$

where σ_{ij} is the conductivity tensor. It is a tensor of rank two because it has two subscripts i and j, each of them runs from 1 to 3. All together it has nine components. The defining property of a tensor is that, when the coordinate axes are rotated, its components must change according to certain transformation rules, analogous to the vector transformation. In fact, a vector, with one subscript attached to its components, is a tensor of rank one. A tensor of nth rank has n subscripts. In this chapter, the mathematics of tensors will be restricted to Cartesian coordinate systems, therefore the name Cartesian tensors.

4.1 Transformation Properties of Vectors

4.1.1 Transformation of Position Vector

The coordinate frame we use to describe positions in space is of course entirely arbitrary, but there is a specific transformation rule for converting vector components from one frame to another. For simplicity, we will first consider a simple case. Suppose the rectangular coordinate system is rotated counterclockwise about the z-axis through an angle θ. The point P is at the position (x, y, z) before the rotation. After the rotation, the same position becomes (x', y', z'), as shown in Fig. 4.1. Therefore the position vector $\mathbf{r}$ expressed in the original system is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \tag{4.1}$$

and expressed in the rotated system is

$$\mathbf{r} = x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}', \tag{4.2}$$

where $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ and $(\mathbf{i}', \mathbf{j}', \mathbf{k}')$ are the unit vectors along the three axes of the original and the rotated coordinates, respectively. The relationship between primed and unprimed systems can be easily found, since

$$\begin{aligned} x' &= \mathbf{i}' \cdot \mathbf{r} = \mathbf{i}' \cdot (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = (\mathbf{i}' \cdot \mathbf{i})\, x + (\mathbf{i}' \cdot \mathbf{j})\, y + (\mathbf{i}' \cdot \mathbf{k}) z \\ &= x\cos\theta + y\cos\left(\frac{\pi}{2} - \theta\right) = x\cos\theta + y\sin\theta \end{aligned} \tag{4.3}$$

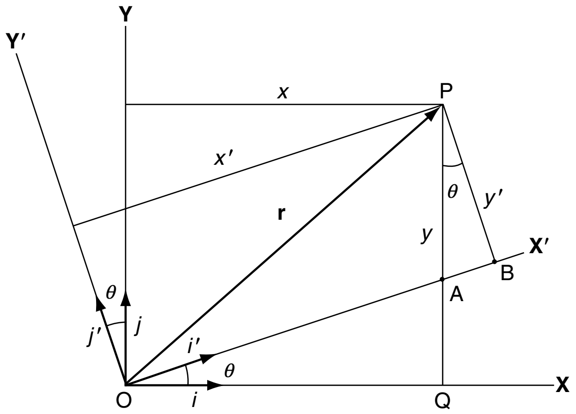

Fig. 4.1. The coordinate system is rotated around z-axis. The primed quantities are those in the rotated system and the unprimed quantities are those in the original system

and

$$\begin{aligned} y' &= \mathbf{j}' \cdot \mathbf{r} = \mathbf{j}' \cdot (x\mathbf{i}+y\mathbf{j}+\mathbf{zk}) = (\mathbf{j}' \cdot \mathbf{i})\, x + (\mathbf{j}' \cdot \mathbf{j})\, y + (\mathbf{i}' \cdot \mathbf{k})z \\ &= x\cos\left(\frac{\pi}{2}+\theta\right) + y\cos\theta = -x\sin\theta + y\cos\theta. \end{aligned} \tag{4.4}$$

Since $\mathbf{k} = \mathbf{k}'$,

$$z' = z.$$

Of course, these relations are also geometrical statements of the rotation. It is seen in Fig. 4.1

$$\begin{aligned} x' &= \mathrm{OA} + \mathrm{AB} = \frac{\mathrm{OQ}}{\cos\theta} + PA\sin\theta = \frac{\mathrm{OQ}}{\cos\theta} + (\mathrm{PQ} - \mathrm{AQ})\sin\theta \\ &= \frac{x}{\cos\theta} + (y - x\tan\theta)\sin\theta = x\left(\frac{1}{\cos\theta} - \frac{\sin^2\theta}{\cos\theta}\right) + y\sin\theta \\ &= x\cos\theta + y\sin\theta, \end{aligned}$$

$$\begin{aligned} y' &= \mathrm{PA}\cos\theta = (\mathrm{PQ} - \mathrm{AQ})\cos\theta = (y - x\tan\theta)\cos\theta \\ &= y\cos\theta - x\sin\theta, \end{aligned}$$

which are identical to (4.3) and (4.4).

With matrix, these relations can be expressed as

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.5}$$

The 3×3 matrix is known as rotation matrix.

4.1.2 Vector Equations

Vector equations are used to express physical laws which should be independent of the reference frame. For example, Newton's second law of dynamics

$$\mathbf{F} = m\mathbf{a}, \tag{4.6}$$

relates the force $\mathbf{F}$ on the particle of mass m and the acceleration $\mathbf{a}$ of the particle. No coordinates appear explicitly in the equation, as it should be since the law is universal. However, often we find it easier to set up a coordinate system and work with the individual components. In any particular coordinate system, each vector is represented by three components. When we change the refrence frame, these components will change. But they must change in a specific way in order for (4.6) to remain true. The coordinates will be changed by either a translation and/or a rotation of the axes. A translation changes the origin of the coordinate system, resulting in some additive constants in the components of $\mathbf{r}$. Since the derivative of a constant is zero, the translation will not affect vectors $\mathbf{F}$ and $\mathbf{a}$. Therefore the important changes are due to the rotation of the axes.

First we note that if (4.6) holds in one coordinate system it holds in all, for the equation may be written as

$$\mathbf{F}{-}m\mathbf{a} = \mathbf{0}, \tag{4.7}$$

and under axes rotation, the zero vector will obviously remain zero in the new system. In terms of its components in Cartesian coordinate system, (4.7) can be written as

$$(F_x - ma_x)\mathbf{i} + (F_y - ma_y)\mathbf{j} + (F_z - ma_z)\mathbf{k} = \mathbf{0}, \tag{4.8}$$

which leads to

$$F_x = ma_x, \quad F_y = ma_y, \quad F_z = ma_z. \tag{4.9}$$

Now if the system is rotated counterclockwise about z-axis through an angle θ as indicated in Fig. 4.1, (4.7) becomes

$$(F'_{x'} - ma'_{x'})\mathbf{i}' + (F'_{y'} - ma'_{y'})\mathbf{j}' + (F'_{z'} - ma'_{z'})\mathbf{k}' = \mathbf{0}, \tag{4.10}$$

where by definition

$$\begin{aligned} a'_{x'} &= \frac{\mathrm{d}^2}{\mathrm{d}t^2}x' = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(x\cos\theta + y\sin\theta) \\ &= a_x\cos\theta + a_y\sin\theta, \end{aligned} \tag{4.11}$$

$$\begin{aligned} a'_{y'} &= \frac{\mathrm{d}^2}{\mathrm{d}t^2}y' = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(-x\sin\theta + y\cos\theta) \\ &= -a_x\sin\theta + a_y\cos\theta, \end{aligned} \tag{4.12}$$

$$a'_{z'} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} z' = \frac{\mathrm{d}^2}{\mathrm{d}t^2} z = a_z. \tag{4.13}$$

Each component of (4.10) must be identically equal to zero. This gives

$$\begin{aligned} F'_{x'} &= ma'_{x'} = m(a_x \cos\theta + a_y \sin\theta), \\ F'_{y'} &= ma'_{y'} = m(-a_x \sin\theta + a_y \cos\theta), \\ F'_{z'} &= ma'_{z'} = ma_z. \end{aligned}$$

Using (4.9), we have

$$\begin{aligned} F'_{x'} &= F_x \cos\theta + F_y \sin\theta, \\ F'_{y'} &= -F_x \sin\theta + F_y \cos\theta, \\ F'_{z'} &= F_z. \end{aligned}$$

Written in the matrix form, these relations are expressed as

$$\begin{pmatrix} F'_{x'} \\ F'_{y'} \\ F'_{z'} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}. \tag{4.14}$$

Comparing (4.5) and (4.14) , we see that the rotation matrix is exactly the same. In other words, the components of the vector $\mathbf{F}$ transform in the same way as those of the position vector $\mathbf{r}$.

In physical applications, it means that in order for a quantity to be considered as a vector, the measured values of its components in a rotated system must be related in this way to those in the original system.

The orientation between two coordinate systems is of course not limited to a single rotation around a particular axis. If we know the relative orientation of the systems, we can follow the procedure of (4.3) to establish the relation

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} (\mathbf{i}' \cdot \mathbf{i}) & (\mathbf{i}' \cdot \mathbf{j}) & (\mathbf{i}' \cdot \mathbf{k}) \\ (\mathbf{j}' \cdot \mathbf{i}) & (\mathbf{j}' \cdot \mathbf{j}) & (\mathbf{j}' \cdot \mathbf{k}) \\ (\mathbf{k}' \cdot \mathbf{i}) & (\mathbf{k}' \cdot \mathbf{j}) & (\mathbf{k}' \cdot \mathbf{k}) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

In Sect. 4.1.3 we consider the actual rotation that will bring $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ into $(\mathbf{i}', \mathbf{j}', \mathbf{k}')$.

4.1.3 Euler Angles

Often we need to express the transformation matrix in terms of concrete rotations which bring the coordinate axes into a specified orientation. In general, the rotation can be regarded as a combination of three rotations, performed successively, about three different directions in space. The most useful description of this kind is in terms of Euler's angles α, β, γ, which we now define.

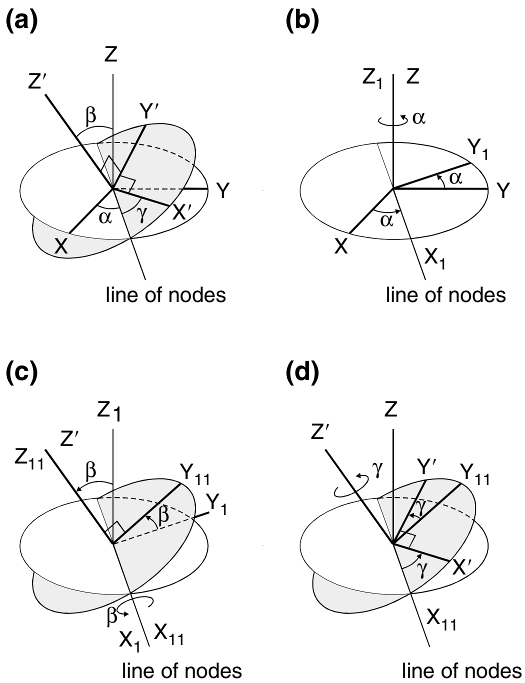

Fig. 4.2. Euler angles. (**a**) Relative orientation of two rectangular coordinate systems XYZ ang $X'Y'Z'$ with a common origin is specified by three Euler angles α, β, γ. The line of nodes is the intersection of XY and X′Y′ planes. The transformation matrix is the product of the matrices representing the following three rotations. (**b**) First rotate α along the Z axis, bring X axis to coincide with the line of nodes. (**c**) Rotate β along the line of nodes. (**d**) Finally rotate γ along Z' axis

The two-coordinate systems are shown in Fig. 4.2a. Let XYZ be the axes of the initial system, $X'Y'Z'$ be the axes of the final system. The intersection of XY plane and X′Y′ plane is known as the line of nodes. The relative orientation of the two systems is specified by the three angles α, β, γ. As shown in Fig. 4.2a, α is the angle between X axis and the line of nodes, β is the angle between Z and Z' axes, and γ is the angle between the line of nodes and X' axis.

The transformation matrix from XYZ to X′Y′Z′ can be obtained by writing it as the product of the separate rotations, each of which has a relative simple rotation matrix. First rotate the initial axes XYZ, by an angle α counterclockwise about the Z axis, bring the X axis to coincide with the line of nodes. The resultant coordinate system is labeled the X_1, Y_1, Z_1 axes, as shown in Fig. 4.2b. In the second stage the intermediate axes are rotated about X_1 axis counterclockwise by an angle β to produce another intermediate X_{11}, Y_{11}, Z_{11} axes, as shown in Fig. 4.2c. Finally the X_{11}, Y_{11}, Z_{11} axes are rotated counterclockwise by an angle γ about the Z_{11} axis to produce the desired $X'Y'Z'$ axes, as shown in Fig. 4.2d.

After the first rotation, the coordinates (x, y, z) of $\mathbf{r}$ in the initial system becomes (x_1, y_1, z_1) in the $X_1\ Y_1\ Z_1$ system. They are related by a rotation matrix,

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.15}$$

The second rotation is about the X_1 axis. After the rotation, (x_1, y_1, z_1) becomes (x_{11}, y_{11}, z_{11}) with the relation

$$\begin{pmatrix} x_{11} \\ y_{11} \\ z_{11} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}. \tag{4.16}$$

After the final rotation about Z_{11} axis, the coordinates of $\mathbf{r}$ becomes (x', y', z') which is given by

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ y_{11} \\ z_{11} \end{pmatrix}. \tag{4.17}$$

It is clear from (4.15) to (4.17) that

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (A) \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{4.18}$$

where

$$(A) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.19}$$

Hence the 3×3 matrix (A) is the rotation matrix of the complete transformation. Multiplying the three matrices out, one can readily find the elements of A,

$$(A) = \begin{pmatrix} \cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha & \cos\gamma\sin\alpha + \sin\gamma\cos\beta\cos\alpha & \sin\gamma\sin\beta \\ -\sin\gamma\cos\alpha - \cos\gamma\cos\beta\sin\alpha & -\sin\gamma\sin\alpha + \cos\gamma\cos\beta\cos\alpha & \cos\gamma\sin\beta \\ \sin\beta\sin\alpha & -\sin\beta\cos\alpha & \cos\beta \end{pmatrix} \tag{4.20}$$

It is not difficult to verify that the product of matrix (A) and its transpose $\left(A^{\mathrm{T}}\right)$ is the identity matrix (I), (The transpose of (A) is the matrix (A) with row and column interchanged.)

$$(A)\left(A^{\mathrm{T}}\right)=(I)\,.$$

Therefore the inverse $(A)^{-1}$ is given by the transpose $(A)^{\mathrm{T}}$,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(A^{\mathrm{T}}\right) \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

These are general properties of a rotation matrix which we shall prove in Sect. 4.1.4

It should be noted that different authors define Euler angles in slightly different ways, because the sequence of rotations used to define the final orientation of the coordinate systems is to some extent arbitrary. We have adopted the definition used in most textbooks on classical mechanics.

4.1.4 Properties of Rotation Matrices

To study the general properties of vector space, it is convenient to use a more systematic notation. Let (x,y,z) be (x_1,x_2,x_3); $(\mathbf{i},\mathbf{j},\mathbf{k})$ be $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$, and (V_x,V_y,V_z) be (V_1,V_2,V_3). The quantities in the rotated system are similarly labeled as prime quantities. One of the advantages of the new notation is that it permits us to use the summation symbol Σ to write the equations in a more compact form. The orthogonality of $(\mathbf{i},\mathbf{j},\mathbf{k})$ is expressed as

$$(\mathbf{e}_i\cdot\mathbf{e}_j)=(\mathbf{e}_i'\cdot\mathbf{e}_j')=\delta_{ij},$$

where the symbol δ_{ij}, known as the Kronecker delta, is defined as

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

In general, the same position vector $\mathbf{r}$, expressed in two different coordinate systems can be written as

$$\mathbf{r}=\sum_{j=1}^{3}x_j'\mathbf{e}_j'=\sum_{j=1}^{3}x_j\mathbf{e}_j. \tag{4.21}$$

Taking the dot product $\mathbf{e}_i'\cdot\mathbf{r}$, we have

$$\mathbf{e}_i'\cdot\sum_{j=1}^{3}x_j'\mathbf{e}_j'=\sum_{j=1}^{3}(\mathbf{e}_i'\cdot\mathbf{e}_j')x_j'=\sum_{j=1}^{3}\delta_{ij}x_j'=x_i'. \tag{4.22}$$

The same dot product from (4.21) gives

$$\mathbf{e}_i' \cdot \sum_{j=1}^{3} x_j \mathbf{e}_j = \sum_{j=1}^{3} (\mathbf{e}_i' \cdot \mathbf{e}_j) x_j. = \sum_{j=1}^{3} a_{ij} x_j. \tag{4.23}$$

It follows from (4.22) and (4.23) that

$$x_i' = \sum_{j=1}^{3} (\mathbf{e}_i' \cdot \mathbf{e}_j) x_j = \sum_{j=1}^{3} a_{ij} x_j, \tag{4.24}$$

where

$$a_{ij} = (\mathbf{e}_i' \cdot \mathbf{e}_j) \tag{4.25}$$

is the direction cosine between $\mathbf{e}_i'$ and $\mathbf{e}_j$. Note that i in (4.24) remains as a parameter which gives rise to three separate equations when it is set to $1, 2,$ and 3. In matrix notation, (4.24) is written as

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{4.26}$$

If, instead of $\mathbf{e}_i' \cdot \mathbf{r}$, we take $\mathbf{e}_i \cdot \mathbf{r}$ and follow the same procedure, we will obtain

$$x_i = \sum_{j=1}^{3} (\mathbf{e}_i \cdot \mathbf{e}_j') x_j'.$$

Since $(\mathbf{e}_i \cdot \mathbf{e}_j')$ is the cosine of the angle between $\mathbf{e}_i$ and $\mathbf{e}_j'$ which can be expressed just as well as $(\mathbf{e}_j' \cdot \mathbf{e}_i)$, and by definition of (4.25) $(\mathbf{e}_j' \cdot \mathbf{e}_i) = a_{ji}$, therefore

$$x_i = \sum_{j=1}^{3} a_{ji} x_j', \tag{4.27}$$

or

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}. \tag{4.28}$$

Comparing (4.26) and (4.28) we see that the inverse of the rotation matrix is equal to its transpose

$$(a_{ij})^{-1} = (a_{ji}) = (a_{ij})^{\mathrm{T}}. \tag{4.29}$$

Any transformation that satisfies this condition is known as an orthogonal transformation.

Renaming the indices i and j, we can write (4.27) as

$$x_j = \sum_{i=1}^{3} a_{ij} x_i'. \tag{4.30}$$

It is thus clear from (4.24) and the last equation that

$$a_{ij} = \frac{\partial x_i'}{\partial x_j} = \frac{\partial x_j}{\partial x_i'}. \tag{4.31}$$

We emphasize this relation is true only in the Cartesian coordinate system.

The nine elements of the rotation matrix are not independent of each other. One way to derive the relationships between them is to note that if the two coordinate systems have the same origin then the length of the position vector should be the same in both systems. This requires

$$\mathbf{r}\cdot\mathbf{r} = \sum_{i=1}^{3} x_i'^2 = \sum_{j=1}^{3} x_j^2. \tag{4.32}$$

Using (4.24) , we have

$$\sum_{i=1}^{3} x_i'^2 = \sum_{i=1}^{3} x_i' x_i' = \sum_{i=1}^{3}\left(\sum_{j=1}^{3} a_{ij}x_j\right)\left(\sum_{k=1}^{3} a_{ik}x_k\right)$$
$$= \sum_{j=1}^{3}\sum_{k=1}^{3} x_j x_k \sum_{i=1}^{3} a_{ij}a_{ik} = \sum_{j=1}^{3} x_j^2.$$

This can be true for all points if and only if

$$\sum_{i=1}^{3} a_{ij}a_{ik} = \delta_{jk}. \tag{4.33}$$

This relation is known as the orthogonality condition. Any matrix whose elements satisfy this condition is called an orthogonal matrix. The rotation matrix is an orthogonal matrix. With all possible values of i and j, (4.33) consists of a set of six equations. This set of equations is equivalent to

$$\sum_{i=1}^{3} a_{ji}a_{ki} = \delta_{jk} \tag{4.34}$$

which can be obtained in the same way from (4.32) , but starting from right to left with the transformation of (4.27) .

Example 4.1.1. Show that the determinant of an orthogonal transformation is equal to either $+1$ or -1.

Solution 4.1.1. Let the matrix of the transformation be (A). Since (A) $(A^{-1}) = (I)$, the determinant of the identity matrix is of course equal to one, $\left|AA^{-1}\right| = 1$. For an orthogonal transformation $A^{-1} = A^{\mathrm{T}}$, so $\left|AA^{\mathrm{T}}\right| = 1$. Since $\left|AA^{\mathrm{T}}\right| = |A|\left|A^{\mathrm{T}}\right|$ and $|A| = \left|A^{\mathrm{T}}\right|$, it follows that $|A|^2 = 1$. Therefore $|A| = \pm 1$.

Example 4.1.2. Show that the determinant of a rotation matrix is equal to +1.

Solution 4.1.2. Express $\mathbf{e}_i'$ in terms of $\{\mathbf{e}_k\}$: $\mathbf{e}_i' = \sum_{i=1}^{3} b_{ik}\mathbf{e}_k$.

$$(\mathbf{e}_i' \cdot \mathbf{e}_j) = \sum_{i=1}^{3} b_{ik}(\mathbf{e}_k \cdot \mathbf{e}_j) = \sum_{i=1}^{3} b_{ik}\delta_{kj} = b_{ij}.$$

But $(\mathbf{e}_i' \cdot \mathbf{e}_j) = a_{ij}$, therefore $b_{ij} = a_{ij}$. So

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

As we have shown in 1.2.7, the scalar triple product is equal to the determinant of the components

$$\mathbf{e}_1' \cdot (\mathbf{e}_2' \times \mathbf{e}_3') = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

which is the rotation matrix. On the other hand,

$$\mathbf{e}_1' \cdot (\mathbf{e}_2' \times \mathbf{e}_3') = \mathbf{e}_1' \cdot \mathbf{e}_1' = +1.$$

Therefore

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +1.$$

4.1.5 Definition of a Scalar and a Vector in Terms of Transformation Properties

Now we come to the refined algebraic definition of a scalar and a vector.

Under a rotation of axes, the coordinates of the position vector in the original system x_i transform to x_i' in the rotated system according to

$$x_i' = \sum_j a_{ij}x_j \tag{4.35}$$

with

$$\sum_i a_{ij}a_{ik} = \delta_{jk}. \tag{4.36}$$

If under such a transformation, a quantity φ is unaffected, then φ is called a scalar. This means if φ is a scalar, then

$$\varphi(x_1, x_2, x_3) = \varphi'(x'_1, x'_2, x'_3). \tag{4.37}$$

Note that after the coordinates are transformed, the functional form may be changed (therefore φ'), but as long as (x_1, x_2, x_3) and (x'_1, x'_2, x'_3) represent the same point, their value is the same.

If a set of quantities (A_1, A_2, A_3) in the original system is transformed into (A'_1, A'_2, A'_3) in the rotated system according to

$$A'_i = \sum_j a_{ij} A_j, \tag{4.38}$$

then the quantity $\mathbf{A} = (A_1, A_2, A_3)$ is called a vector. Since $(a_{ij})^{-1} = (a_{ji})$, (4.38) is equivalent to

$$A_i = \sum_j a_{ji} A'_j. \tag{4.39}$$

This definition is capable of generalization and ensures that vector equations are independent of coordinate system.

Example 4.1.3. Suppose $\mathbf{A}$ and $\mathbf{B}$ are vectors. Show that the dot product $\mathbf{A} \cdot \mathbf{B}$ is a scalar.

Solution 4.1.3. Since $\mathbf{A}$ and $\mathbf{B}$ are vectors, under a rotation their components transform according to

$$A'_i = \sum_j a_{ij} A_j; \quad B'_i = \sum_j a_{ij} B_j.$$

To show the dot product

$$\mathbf{A} \cdot \mathbf{B} = \sum_i A_i B_i$$

is a scalar, we must show that its value in the rotated system is the same as its value in the original system.

$$(\mathbf{A} \cdot \mathbf{B})' = \sum_i A'_i B'_i = \sum_i \left(\sum_j a_{ij} A_j \right) \left(\sum_k a_{ik} A_k \right)$$

$$= \sum_j \sum_k \left(\sum_i a_{ij} a_{ik} \right) A_j A_k = \sum_j \sum_k \delta_{jk} A_j B_k = \sum_j A_j B_j = \mathbf{A} \cdot \mathbf{B}.$$

So $\mathbf{A} \cdot \mathbf{B}$ is a scalar.

Example 4.1.4. Show that if (A_1, A_2, A_3) is such that $\sum_i A_i B_i$ is a scalar for every vector $\mathbf{B}$, then (A_1, A_2, A_3) is a vector.

Solution 4.1.4. Since $\sum_i A_i B_i$ is a scalar and $\mathbf{B}$ is a vector,

$$\sum_i A_i B_i = \sum_i A_i' B_i' = \sum_i A_i' \sum_j a_{ij} B_j.$$

Now both i and j are running indices, we can rename i as j, and j as i. So

$$\sum_i A_i B_i = \sum_j A_j' \sum_i a_{ji} B_i = \sum_i \sum_j a_{ji} A_j' B_i.$$

It follows

$$\sum_i \left(A_i - \sum_j a_{ji} A_j' \right) B_i = 0.$$

Since this identity holds for every $\mathbf{B}$, we must have

$$A_i = \sum_j a_{ji} A_j'.$$

Therefore A_1, A_2, A_3 are components of a vector.

Example 4.1.5. Show that, in Cartesian coordinates, the gradient of a scalar function $\boldsymbol{\nabla}\varphi$ is a vector function.

Solution 4.1.5. As a scalar it must have the same value at a given point in space, independent of the orientation of the coordinate system

$$\varphi'(x_1', x_2', x_3') = \varphi(x_1, x_2, x_3). \tag{4.40}$$

Differentiating with respect to x_i' and using the chain rule, we have

$$\frac{\partial}{\partial x_i'} \varphi'(x_1', x_2', x_3') = \frac{\partial}{\partial x_i'} \varphi(x_1, x_2, x_3) = \sum_j \frac{\partial \varphi}{\partial x_j} \frac{\partial x_j}{\partial x_i'}. \tag{4.41}$$

It follows from (4.31) that in Cartesian coordinates

$$\frac{\partial x_j}{\partial x_i'} = a_{ij},$$

therefore

$$\frac{\partial \varphi'}{\partial x_i'} = \sum_j a_{ij} \frac{\partial \varphi}{\partial x_j}. \tag{4.42}$$

Now the components of $\nabla\varphi$ are

$$\left(\frac{\partial\varphi}{\partial x_1}, \frac{\partial\varphi}{\partial x_2}, \frac{\partial\varphi}{\partial x_3}\right).$$

They transform under a rotation of coordinates in exactly the same way as the components of a vector, therefore $\nabla\varphi$ is a vector function.

A vector whose components are just numbers is a constant vector. All constant vectors behave like a position vector. When the axes are rotated, the components change into a new set of numbers in accordance with the transformation rule. Therefore, any set of three numbers can be considered as a constant vector.

For vector fields, the components are functions of (x_1, x_2, x_3) themselves. Under a rotation, not only (x_1, x_2, x_3) will change to (x_1', x_2', x_3'), the appearances of the component functions may also change. This leads to some complications.

Mathematically the transformation rules place little restriction on what we can call a vector. We can make any set of three functions the components of a vector field by simply defining, in a rotated system, the corresponding functions obtained from the correct transformation rules as the components of the vector in that system.

However, if we are discussing a physical entity, we are not free to define its components in various systems. They are determined by physical facts. As we stated earlier, all properly formulated physical laws must be independent of the coordinate system. In other words, the appearance of the equation describing physical laws must be the same in all coordinate systems. If vector functions maintain the same appearances in rotated systems, equations written in terms of them will be automatically invariant under rotation. Therefore we include in the definition of a vector field, an additional condition that the transformed components must look the same as the original components.

For example, many authors describe

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \tag{4.43}$$

as a vector field in a two-dimensional space (see, for example, E.M. Purcell, "Electricity and Magnetism", McGraw-Hill Book Co. (1965), page 36; D.A. McQuarrie, "Mathematical Methods for Scientists and Engineers," University Science Books, (2003), page 301), still many others would say that (4.43) cannot be called a vector field (see, for example, G. Arfken, "Mathematical Methods for Physicists", Academic Press, (1968), page 8; P.C. Mathews, "Vector Calculus", Springer, Berlin Heidelberg New York (2002), page 118).

If we consider (4.43) as a vector, then the components of this vector in the system where the axes are rotated by an angle θ, are given by

$$\begin{pmatrix} V_1' \\ V_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

Furthermore, the coordinates have to change according to

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

One can readily show that

$$\begin{pmatrix} V_1' \\ V_2' \end{pmatrix} = \begin{pmatrix} 2x_1'\sin\theta\cos\theta + x_2'(\cos^2\theta - \sin^2\theta) \\ x_1'(\cos^2\theta - \sin^2\theta) - 2x_2'\sin\theta\cos\theta \end{pmatrix}. \tag{4.44}$$

Mathematically one can certainly define (4.44) as the components of the vector in the rotated system, but they do not look like (4.43) .

On the other hand, consider a slightly different expression

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}. \tag{4.45}$$

With the same transformation rules, we obtain

$$\begin{pmatrix} V_1' \\ V_2' \end{pmatrix} = \begin{pmatrix} x_2'(\cos^2\theta + \sin^2\theta) \\ -x_1'(\cos^2\theta + \sin^2\theta) \end{pmatrix} = \begin{pmatrix} x_2' \\ -x_1' \end{pmatrix}, \tag{4.46}$$

which has the same form as (4.45) . In this sense, we say that (4.45) is invariant under rotations

Under our definition, (4.45) is a vector and (4.43) is not.

4.2 Cartesian Tensors

4.2.1 Definition

The definition of a vector can be extended to define a more general class of objects called tensors, which may have more than one subscript.

If in the rectangular coordinate system of three-dimensional space, under a rotation of coordinates

$$x_i' = \sum_{j=1}^{3} a_{ij} x_j,$$

the 3^N quantities $T_{i_1,i_2,\cdots,i_N}$ (where each of $i_1, i_2, \cdots, i_N$ runs independently from 1 to 3) transform according to the rule

$$T'_{i_1,i_2,\cdots,i_N} = \sum_{j_1=1}^{3}\sum_{j_2=1}^{3}\cdots\sum_{j_N=1}^{3} a_{i_1j_1}a_{i_2j_2}\cdots a_{i_Nj_N}T_{j_1,j_2,\cdots,j_N}, \tag{4.47}$$

then $T_{i_1,i_2,\cdots,i_N}$ are the components of a Nth rank Cartesian tensor. Since we are going to restrict our discussion to Cartesian tensors, unless explicitly otherwise specified, we shall drop the word Cartesian from here on.

The rank of a tensor is the number of free subscripts. Tensor of zeroth rank has only one ($3^0 = 1$) component. So it can be thought as a scalar. A tensor of first rank has three components ($3^1 = 3$). The rule of transformation of these components under a rotation is the same as the rule for a vector. So a vector is a tensor of rank one.

The most useful other case is the tensor of second rank. It has nine components ($3^2 = 9$), T_{ij} which obey the transformation rule

$$T'_{ij} = \sum_{l=1}^{3}\sum_{m=1}^{3} a_{il}a_{jm}T_{lm}. \tag{4.48}$$

The components of a second rank tensor may be conveniently expressed as a 3×3 matrix:

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

However, this does not mean that any 3×3 matrix forms a tensor. The essential condition is that its components satisfy the transformation rule.

As a matter of terminology, a second-rank tensor in three-dimensional space is a collection of nine components T_{ij}. However, very often T_{ij} is referred to as "tensor" instead of "tensor components" for simplicity. In other words, T_{ij} is used to mean the totality of the components as well as the individual component. The context will make its meaning clear. Another often used symbol for tensors is a double bar over a letter, such as $\overline{\overline{T}}$.

Example 4.2.1. Show that in a two-dimensional space, the following quantity is a second-rank tensor

$$T_{ij} = \begin{pmatrix} x_1x_2 & -x_1^2 \\ x_2^2 & -x_1x_2 \end{pmatrix}.$$

Solution 4.2.1. In a two-dimensional space, a second rank tensor has four ($2^2 = 4$) components. If it is a tensor, in a rotated system it must look like

$$T'_{ij} = \begin{pmatrix} x'_1x'_2 & -x'^2_1 \\ x'^2_2 & -x'_1x'_2 \end{pmatrix},$$

where

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Now we must check if each component satisfies the transformation rule.

$$\begin{aligned} T_{11}' &= x_1' x_2' = (\cos\theta x_1 + \sin\theta x_2)(-\sin\theta x_1 + \cos\theta x_2) \\ &= -\cos\theta\sin\theta x_1^2 + \cos^2\theta x_1 x_2 - \sin^2\theta x_2 x_1 + \sin\theta\cos\theta x_2^2. \end{aligned}$$

This is to be compared with

$$\begin{aligned} T_{11}' &= \sum_{l=1}^{2}\sum_{m=1}^{2} a_{1l} a_{1m} T_{lm} \\ &= a_{11}a_{11}T_{11} + a_{11}a_{12}T_{12} + a_{12}a_{11}T_{21} + a_{12}a_{12}T_{22} \\ &= \cos^2\theta x_1 x_2 - \cos\theta\sin\theta x_1^2 + \sin\theta\cos\theta x_2^2 - \sin^2\theta x_1 x_2. \end{aligned}$$

It is seen that these two expressions are identical. The same process will show that other components will also satisfy the transformation rule. Therefore T_{ij} is a second rank tensor in the two-dimensional space.

This transformation property is not to be taken for granted. In the above example, if one algebraic sign is changed, the transformation rule will not be satisfied. For example if T_{22} is changed to $x_1 x_2$,

$$T_{ij} = \begin{pmatrix} x_1 x_2 & -x_1^2 \\ x_2^2 & x_1 x_2 \end{pmatrix}, \tag{4.49}$$

then

$$T_{11}' \neq \sum_{l=1}^{2}\sum_{m=1}^{2} a_{1l} a_{1m} T_{lm}.$$

Therefore (4.49) is not a tensor.

4.2.2 Kronecker and Levi-Civita Tensors

Kronecker delta tensor. The Kronecker delta which we have already encountered,

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

is a second rank tensor. To prove this, consider the transformation

$$\delta_{ij}' = \sum_{l=1}^{3}\sum_{m=1}^{3} a_{il} a_{jm} \delta_{lm} = \sum_{l=1}^{3} a_{il} a_{jl} = \delta_{ij}. \tag{4.50}$$

Therefore

$$\delta'_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Thus δ_{ij} obeys the tensor transformation rule and is invariant under rotation. In addition, it has a special property. The numerical values of its components are the same in all coordinate systems. A tensor with this property is known as an isotropic tensor.

Since

$$\sum_k D_{ik}\delta_{jk} = D_{ij},$$

the Kronecker delta tensor is also known as the *substitution tensor*. It is also called a unit tensor, because of its matrix representation:

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Levi-Civita Tensor. The Levi-Civita symbol ε_{ijk},

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 & \text{if any index is repeated} \end{cases},$$

which we used for the definition of a third-order determinant, is a third rank isotropic tensor. This is also known as the *alternating tensor*. To prove this, recall the definition of the third-order determinant

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{1l}a_{2m}a_{3n}\varepsilon_{lmn} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{4.51}$$

Now if the row indices (1, 2, 3) are replaced by (i, j, k), we have

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn} = \begin{vmatrix} a_{i1} & a_{i2} & a_{i3} \\ a_{j1} & a_{j2} & a_{j3} \\ a_{k1} & a_{k2} & a_{k3} \end{vmatrix}. \tag{4.52}$$

This relation can be demonstrated by writing out the nonvanishing terms of both sides. It can also be proved by noting the following. First for $i = 1$, $j = 2$, $k = 3$, it reduces to (4.51). Now consider the effect of interchanging i and j. The left-hand side changes sign because

$$\begin{aligned} \sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{jl}a_{im}a_{kn}\varepsilon_{lmn} &= \sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{jm}a_{il}a_{kn}\varepsilon_{m\ln} \\ &= -\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn}. \end{aligned}$$

The right-hand side also changes sign because it is an interchange of two rows of the determinant. If two indices of i, j, k are the same, then both sides are equal to zero. The left-hand side is zero, since the quantity is equal to its negative. The right-hand side is equal to zero since two rows of the determinant are identical. This suffices to prove the result since all permutations of i, j, k can be achieved by a sequence of interchanges.

It follows from the properties of determinants and the definition of ε_{ijk} that

$$\begin{vmatrix} a_{i1} & a_{i2} & a_{i3} \\ a_{j1} & a_{j2} & a_{j3} \\ a_{k1} & a_{k2} & a_{k3} \end{vmatrix} = \varepsilon_{ijk} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \tag{4.53}$$

and (4.52) becomes

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn} = \varepsilon_{ijk} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{4.54}$$

This relation is true for any determinant. Now if a_{ij} are elements of a rotation matrix,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 1,$$

as shown in example 4.1.2.

To decide if ε_{ijk} is a tensor, we should look at its value in a rotated system. The tensor transformation rules require

$$\varepsilon'_{ijk} = \sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn}.$$

But

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn} = \varepsilon_{ijk} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \varepsilon_{ijk}.$$

Hence

$$\varepsilon'_{ijk} = \varepsilon_{ijk}. \tag{4.55}$$

Therefore, ε_{ijk} is indeed a third-rank isotropic tensor.

Relation between δ_{ij} and ε_{ijk}. There is an interesting and important relation between Kronecker delta and Levi-Civita tensors

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{4.56}$$

After summed over i, there are four free subscripts j, k, l, m. Therefore (4.56) represents 81 ($3^4 = 81$) equations. Yet it is not difficult to prove (4.56), when we make the following observations.

1. If either $j = k$ or $l = m$, both sides of (4.56) are equal to zero. If $j = k$, the left-hand side is zero, because $\varepsilon_{ikk} = 0$. The right-hand side is also equal to zero because $\delta_{kl}\delta_{km} - \delta_{km}\delta_{kl} = 0$. The same result is obtained for $l = m$. Therefore we only need to check the cases for which $j \neq k$ and $l \neq m$.
2. For the left-hand side not to vanish, $i,\ j,\ k$ have to be different. Therefore given $j \neq k,\ i$ is fixed. Consider ε_{ilm}, since i is fixed and $i,\ l,\ m$ have to be different for nonvanishing ε_{ilm}, therefore, $l = j,\ m = k$ or $l = k,\ m = j$ are the only two nonvanishing options.
3. For $l = j$ and $m = k$, $\varepsilon_{ijk} = \varepsilon_{ilm}$. Thus ε_{ijk} and ε_{ilm} must have the same sign. (Either both equal to -1, or both equal to $+1$). Therefore on the left-hand side of (4.56), $\varepsilon_{ijk}\varepsilon_{ilm} = +1$. On the right-hand side of (4.56), it is also equal to $+1$, since $l \neq m$,
$$\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} = \delta_{ll}\delta_{mm} - \delta_{lm}\delta_{ml} = 1 - 0 = 1.$$
4. For $l = k$ and $m = j$, $\varepsilon_{ijk} = \varepsilon_{iml} = -\varepsilon_{ilm}$. Thus ε_{ijk} and ε_{ilm} have opposite sign. (one equal to -1 and the other $+1$, or vice versa). Therefore on the left-hand side of (4.56), $\varepsilon_{ijk}\varepsilon_{ilm} = -1$. On the right-hand side of (4.56), it is also equal to -1, since $l \neq m$,
$$\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} = \delta_{ml}\delta_{lm} - \delta_{mm}\delta_{ll} = 0 - 1 = -1.$$

This covers all 81 cases. In each case the left-hand side is equal to the right-hand side. Therefore (4.56) is established.

4.2.3 Outer Product

If $S_{i_1 i_2 \cdots i_N}$ *is a tensor of rank* N *and* $T_{j_1 j_2 \cdots j_M}$ *is a tensor of* Mth rank, *then* $S_{i_1 i_2 \cdots i_N} T_{j_1 j_2 \cdots j_M}$ *is a tensor of rank* $(N + M)$.

This is known as the *outer product* theorem. (Outer product is also known as direct product.) This theorem can be easily demonstrated. First it certainly has 3^{N+M} components. Under a rotation
$$\begin{aligned} S'_{i_1 i_2 \cdots i_N} &= \sum_{k_1 \cdots k_N} a_{i_1 k_1} \cdots a_{i_N k_N} S_{k_1 \cdots k_N}, \\ T'_{j_1 j_2 \cdots j_M} &= \sum_{l_1 \cdots l_N} a_{j_1 l_1} \cdots a_{j_M l_M} T_{l_1 \cdots l_M}, \end{aligned}$$
where we have written $\sum_{j_1=1}^{3} \sum_{j_2=1}^{3} \cdots \cdot \sum_{j_N=1}^{3}$ as $\sum_{j_1 \cdots j_N}$,
$$\begin{aligned} &(S_{i_1 i_2 \cdots i_N} T_{j_1 j_2 \cdots j_M})' = S'_{i_1 i_2 \cdots i_N} T'_{j_1 j_2 \cdots j_M} \\ &= \sum_{k_1 \cdots k_N} \sum_{l_1 \cdots l_N} a_{i_1 k_1} \cdots a_{i_N k_N} a_{j_1 l_1} \cdots a_{j_M l_M} S_{k_1 \cdots k_N} T_{l_1 \cdots l_M}, \end{aligned} \tag{4.57}$$
which is how a $(M + N)$th rank tensor should transform.

For example, the outer product of two vectors is a second rank tensor. Let (A_1, A_2, A_3) and (B_1, B_2, B_3) be vectors, so they are first rank tensors. Their outer product A_iB_j is a second rank tensor. Its nine components can be displayed as a matrix

$$A_iB_j = \begin{pmatrix} A_1B_1 & A_1B_2 & A_1B_3 \\ A_2B_1 & A_2B_2 & A_2B_3 \\ A_3B_1 & A_3B_2 & A_3B_3 \end{pmatrix}.$$

Since **A** and **B** are vectors,

$$A_i' = \sum_{k=1}^{3} a_{ik}A_k; \quad B_j' = \sum_{l=1}^{3} a_{jl}B_l,$$

it follows

$$A_i'B_j' = \sum_{k=1}^{3}\sum_{l=1}^{3} a_{ik}a_{jl}A_kB_l,$$

which shows A_iB_j is a second rank tensor, in agreement with the outer product theorem.

We mention in passing, the second rank tensor formed by two vectors **A** and **B** is sometimes denoted as **AB** (without anything between them). When written in this way, it is called a dyad. A linear combination of dyads is a dyadic. Since everything that can be done with vectors and dyadics can also be done by tensors and matrices, but not the other way around, we will not discuss dyadics any further.

Example 4.2.2. Use the outer product theorem to show that the expression in example 4.2.1

$$T_{ij} = \begin{pmatrix} x_1x_2 & -x_1^2 \\ x_2^2 & -x_1x_2 \end{pmatrix}$$

is a second rank tensor in a two-dimensional space.

Solution 4.2.2. A two-dimensional position vector is given by

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We have also shown in (4.46) that

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}$$

is a vector in the two-dimensional space. The outer product of these two vectors is a second rank tensor

$$A_iB_j = \begin{pmatrix} x_1x_2 & -x_1^2 \\ x_2^2 & -x_1x_2 \end{pmatrix} = T_{ij}.$$

4.2.4 Contraction

We can lower the rank of any tensor through the following theorem.

If $T_{i_1 i_2 i_3 \cdots i_N}$ *is a tensor of* Nth *rank, then*

$$S_{i_3 \cdots i_N} = \sum_{i_1} \sum_{i_2} \delta_{i_1 i_2} T_{i_1 i_2 i_3 \cdots i_N}$$

is a tensor of rank $(N-2)$.

To prove this theorem, we first note that $S_{i_3 \cdots i_N}$ has 3^{N-2} components. Next we have to show that

$$S'_{i_3 \cdots i_N} = \sum_{i_1} \sum_{i_2} \delta'_{i_1 i_2} T'_{i_1 i_2 i_3 \cdots i_N} \tag{4.58}$$

satisfies the tensor transformation rule.

With

$$\begin{aligned} \delta'_{i_1 i_2} &= \delta_{i_1 i_2}, \\ T'_{i_1, i_2, \cdots i_N} &= \sum_{j_1 \cdots j_N} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_N j_N} T_{j_1, j_2, \cdots, j_N}, \end{aligned}$$

(4.58) becomes

$$\begin{aligned} S'_{i_3 \cdots i_N} &= \sum_{i_1 i_2} \delta_{i_1 i_2} \sum_{j_1 \cdots j_N} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_N j_N} T_{j_1, j_2, \cdots, j_N} \\ &= \sum_{j_1 \cdots j_N} \left(\sum_{i_1 i_2} \delta_{i_1 i_2} a_{i_1 j_1} a_{i_2 j_2} \right) a_{i_3 j_3} \cdots a_{i_N j_N} T_{j_1, j_2, \cdots, j_N}. \end{aligned}$$

Now

$$\sum_{i_1 i_2} \delta_{i_1 i_2} a_{i_1 j_1} a_{i_2 j_2} = \sum_{i_1} a_{i_1 j_1} a_{i_1 j_2} = \delta_{j_1 j_2},$$

so

$$\begin{aligned} S'_{i_3 \cdots i_N} &= \sum_{j_1 \cdots j_N} \delta_{j_1 j_2} a_{i_3 j_3} \cdots a_{i_N j_N} T_{j_1, j_2, \cdots, j_N} \\ &= \sum_{j_3 \cdots j_N} \left(\sum_{j_1 j_2} \delta_{j_1 j_2} T_{j_1, j_2, \cdots, j_N} \right) a_{i_3 j_3} \cdots a_{i_N j_N} \\ &= \sum_{j_3 \cdots j_N} a_{i_3 j_3} \cdots a_{i_N j_N} S_{j_3 \cdots j_N}. \end{aligned} \tag{4.59}$$

Therefore $S_{i_3 \cdots i_N}$ is a $(N-2)$th rank tensor.

The process of multiplying by $\delta_{i_1 i_2}$ and summing over i_1 and i_2 is called *contraction*. For example, we have shown that $A_i B_j$, the outer product of two vectors **A** and **B**, is a second rank tensor. The contraction of this second-rank tensor is a zeroth rank tensor, namely a scalar.

$$\sum_{ij} \delta_{ij} A_i B_j = \sum_i A_i B_i = \mathbf{A} \cdot \mathbf{B}.$$

Indeed, this zeroth rank tensor is the dot product of **A** and **B**.

Simply stated, a new tensor of rank $(N-2)$ will be obtained if two of the indices of a Nth rank tensor are set equal to each other and summed over. (The German word for contraction is *verjüngung,* which can be translated as rejuvenation.) If the rank of the tensor is 3 or higher, we can contract over any two indices. In general we get different $(N-2)$th rank tensors if we contract over different pairs of indices. For example, in the third rank tensor $T_{i_1 i_2 i_3} = A_{i_1} B_{i_2} C_{i_3}$, if i_1 and i_2 are contracted, we obtain a vector

$$\sum_i T_{iii_3} = \sum_i A_i B_i C_{i_3} = (\mathbf{A} \cdot \mathbf{B})\, C_{i_3},$$

(remember C_{i_3} could represent a particular component of **C**, it also could represent the totality of the components, namely the vector **C** itself). On the other hand, if i_2 and i_3 are contracted, we obtain another vector

$$\sum_i T_{i_1 ii} = \sum_i A_{i_1} B_i C_i = A_{i_1} (\mathbf{B} \cdot \mathbf{C})\,.$$

So contracting the first two indices we get a scalar times vector **C**, and contracting the second and third indices we get another scalar times vector **A**.

Contraction is one of the most important operations in tensor analysis and is worth remembering.

4.2.5 Summation Convention

The summation convention (invented by Einstein) gives tensor analysis much of its appeal. We note that in the definition of tensors (4.47), all indices over which the summation is to be carried out are repeated in the expression. Moreover, the range of the index (such as from 1 to 3) is already known from the context of the discussion. Therefore, without loss of information, we may drop the summation sign with the understanding that the repeated subscripts imply that the term is to be summed over its range. The repeated subscript is referred to as a “dummy” subscript. It must appear no more than twice in a term. The choice of dummy subscript does not matter. Replace one dummy index by another is one of the most useful tricks in tensor analysis one should learn. For example, the dot product $\mathbf{A} \cdot \mathbf{B}$ can be equally represented either by $A_i B_i$ or $A_k B_k$, since both of them mean the same thing, namely

$$A_k B_k = \sum_{i=1}^{3} A_i B_i = \sum_{j=1}^{3} A_j B_j = A_1 B_1 + A_2 B_2 + A_3 B_3 = \mathbf{A} \cdot \mathbf{B}\,.$$

Example 4.2.3. Express the expression $A_i B_j C_i$ with summation convention in terms of ordinary vector notation.

Solution 4.2.3.

$$A_i B_j C_i = \left(\sum_{i=1}^{3} A_i C_i\right) B_j = (\mathbf{A} \cdot \mathbf{C})\, B_j.$$

Note that it is the subscripts that indicate which vector is dotted with which, not the ordering of the components of the vectors. The ordering is immaterial. So $(\mathbf{A} \cdot \mathbf{C})\, B_j = A_k C_k B_j$ is equally valid. The letter j is a free subscript, it can be replaced by any letter other than the dummy subscript. However, if the term is used in a equation, then the free subscript of every term in the equation must be represented by the same letter.

From now on when we write down a quantity with N subscripts, if all N subscripts are different, it will be assumed that it is a good Nth rank tensor. If any two of them are the same, it is a contracted tensor of rank $N-2$.

Example 4.2.4. (a) What is the rank of the tensor $\varepsilon_{ijk} A_l B_m$? (b) what is the rank of the tensor $\varepsilon_{ijk} A_j B_k$? (c) Express $\varepsilon_{ijk} A_j B_k$ in terms of ordinary vector notation.

Solution 4.2.4. (a) Since ε_{ijk} is a third rank tensor and $A_l B_m$ is a tensor of rank 2, so by the outer product theorem $\varepsilon_{ijk} A_l B_m$ is a tensor of rank 5. (b) $\varepsilon_{ijk} A_j B_k$ is twice contracted, so it is a first rank $(5-4=1)$ tensor. (c)

$$\varepsilon_{ijk} A_j B_k = \sum_j \sum_k \varepsilon_{ijk} A_j B_k.$$

If $i=1$, the only nonzero terms come from $j = 2$ *or* 3, since ε_{ijk} is equal to zero if any two indices are equal. If $j=2$, k can only equal to 3. If $j=3$, k has to be 2. So

$$\varepsilon_{1jk} A_j B_k = \varepsilon_{123} A_2 B_3 + \varepsilon_{132} A_3 B_2 = A_2 B_3 - A_3 B_2 = (\mathbf{A} \times \mathbf{B})_1.$$

Similarly,

$$\varepsilon_{2jk} A_j B_k = (\mathbf{A} \times \mathbf{B})_2, \qquad \varepsilon_{3jk} A_j B_k = (\mathbf{A} \times \mathbf{B})_3.$$

Therefore

$$\varepsilon_{ijk} A_j B_k = (\mathbf{A} \times \mathbf{B})_i.$$

Example 4.2.5. Show that

$$\varepsilon_{ijk} A_i B_j C_k = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix}.$$

Solution 4.2.5.

$$\begin{aligned} \varepsilon_{ijk} A_i B_j C_k &= (\varepsilon_{ijk} A_i B_j) C_k = (\mathbf{A} \times \mathbf{B})_k C_k = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} \\ &= \begin{vmatrix} C_1 & C_2 & C_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix}. \end{aligned}$$

With the summation convention, (4.56) is simply

$$\boxed{\varepsilon_{ijk} \varepsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}.} \tag{4.60}$$

Many vector identities can be quickly and elegantly proved with this equation.

Example 4.2.6. Show that

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - \mathbf{C} (\mathbf{A} \cdot \mathbf{B})$$

Solution 4.2.6.

$$\begin{aligned} (\mathbf{B} \times \mathbf{C})_i &= \varepsilon_{ijk} B_j C_k \\ [\mathbf{A} \times (\mathbf{B} \times \mathbf{C})]_l &= \varepsilon_{lmn} A_m (\mathbf{B} \times \mathbf{C})_n = \varepsilon_{lmn} A_m \varepsilon_{njk} B_j C_k = \varepsilon_{nlm} \varepsilon_{njk} A_m B_j C_k \end{aligned}$$

$$\begin{aligned} \varepsilon_{nlm} \varepsilon_{njk} A_m B_j C_k &= (\delta_{lj} \delta_{mk} - \delta_{lk} \delta_{mj}) A_m B_j C_k \\ &= A_k B_l C_k - A_j B_j C_l = (\mathbf{A} \cdot \mathbf{C}) B_l - (\mathbf{A} \cdot \mathbf{B}) C_l. \end{aligned}$$

Since the corresponding components agree, the identity is established.

Example 4.2.7. Show that

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C}) (\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D}) (\mathbf{B} \cdot \mathbf{C}).$$

Solution 4.2.7.

$$\begin{aligned} (\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= \varepsilon_{kij} A_i B_j \varepsilon_{klm} C_l D_m \\ = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) A_i B_j C_l D_m &= A_l B_m C_l D_m - A_m B_l C_l D_m \\ &= (\mathbf{A} \cdot \mathbf{C}) (\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D}) (\mathbf{B} \cdot \mathbf{C}). \end{aligned}$$

4.2.6 Tensor Fields

A tensor field of Nth rank, $T_{i_1 \cdots i_N}(x_1, x_2, x_3)$ is the totality of 3^N functions which for any given point in space (x_1, x_2, x_3) constitute a tensor of Nth rank.

A scalar field is a tensor field of rank zero. We have shown in example 4.1.5 that the gradient of a scalar field is a vector field. There is a corresponding theorem for tensor fields.

If $T_{i_1\cdots i_N}(x_1, x_2, x_3)$ *is a tensor field of rank* N, then

$$\frac{\partial}{\partial x_i} T_{i_1\cdots i_N}(x_1, x_2, x_3)$$

is a tensor field of rank $N+1$.

The proof of this theorem goes as follows.

$$\begin{aligned}\left(\frac{\partial}{\partial x_i} T_{i_1\cdots i_N}(x_1, x_2, x_3)\right)' &= \frac{\partial}{\partial x_i'} T'_{i_1\cdots i_N}(x_1', x_2', x_3') \\ &= \frac{\partial}{\partial x_i'} \sum_{j_1\cdots j_N} a_{i_1 j_1} \cdots a_{i_N j_N} T_{j_1\cdots j_N}(x_1, x_2, x_3).\end{aligned}$$

By chain rule and (4.31)

$$\frac{\partial}{\partial x_i'} = \sum_j \frac{\partial x_j}{\partial x_i'} \frac{\partial}{\partial x_j} = \sum_j a_{ij} \frac{\partial}{\partial x_j},$$

so

$$\left(\frac{\partial}{\partial x_i} T_{i_1\cdots i_N}(x_1, x_2, x_3)\right)' = \sum_j \sum_{j_1\cdots j_N} a_{ij} a_{i_1 j_1} \cdots a_{i_N j_N} \frac{\partial}{\partial x_j} T_{j_1\cdots j_N}(x_1, x_2, x_3), \tag{4.61}$$

which is how a tensor of rank $N+1$ transforms. Therefore the theorem is proven.

To simplify the writing, we introduce another useful notation. From now on, the differential operator $\partial/\partial x_i$ is denoted by ∂_i. For example,

$$\frac{\partial}{\partial x_i}\varphi = \partial_i \varphi,$$

which is the ith component of $\boldsymbol{\nabla}\varphi$. Again it can also represent the totality of

$$\left(\frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2}, \frac{\partial \varphi}{\partial x_3}\right).$$

Thus $\partial_i \varphi$ is a vector. (Note it has one subscript.)

Similarly,

$$\boldsymbol{\nabla} \times \mathbf{A} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ A_1 & A_2 & A_3 \end{vmatrix},$$

$$\boldsymbol{\nabla} \cdot \mathbf{A} = \sum_i \frac{\partial A_i}{\partial x_i} = \sum_i \partial_i A_i = \partial_i A_i,$$

$$(\boldsymbol{\nabla} \times \mathbf{A})_i = \sum_{j,k} \varepsilon_{ijk} \frac{\partial}{\partial x_j} A_k = \varepsilon_{ijk} \partial_j A_k$$

$$\nabla^2 \varphi = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \varphi = \partial_i (\boldsymbol{\nabla} \varphi)_i = \partial_i \partial_i \varphi.$$

With these notations, vector field identities can be easily established.

Example 4.2.8. Show that $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = 0$.

Solution 4.2.8.

$$\begin{aligned}
\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) &= \partial_i \left(\boldsymbol{\nabla} \times \mathbf{A}\right)_i = \partial_i \varepsilon_{ijk} \partial_j A_k = \varepsilon_{ijk} \partial_i \partial_j A_k \\
&= \varepsilon_{ijk} \partial_j \partial_i A_k \qquad (\partial_i \partial_j = \partial_j \partial_i) \\
&= -\varepsilon_{jik} \partial_j \partial_i A_k \\
&= -\varepsilon_{ijk} \partial_i \partial_j A_k \qquad (\text{rename } \; i \text{ and } j) \\
&= -\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A})
\end{aligned}$$

Thus $2\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = 0$. Hence $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = 0$.

Example 4.2.9. Show that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \varphi) = 0$.

Solution 4.2.9.

$$\begin{aligned}
\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \varphi\right]_i &= \varepsilon_{ijk} \partial_j \left(\boldsymbol{\nabla} \varphi\right)_k = \varepsilon_{ijk} \partial_j \partial_k \varphi = -\varepsilon_{ikj} \partial_j \partial_k \varphi \\
&= -\varepsilon_{ijk} \partial_j \partial_k \varphi \qquad (\text{rename } j \text{ and } k) \\
&= -\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \varphi\right]_i .
\end{aligned}$$

It follows that $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \varphi = 0$.

Example 4.2.10. Show that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \mathbf{A}\right) - \nabla^2 \mathbf{A}$.

Solution 4.2.10.

$$\begin{aligned}
\left[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A})\right]_i &= \varepsilon_{ijk} \partial_j \left(\boldsymbol{\nabla} \times \mathbf{A}\right)_k = \varepsilon_{ijk} \partial_j \varepsilon_{klm} \partial_l A_m \\
&= \varepsilon_{kij} \varepsilon_{klm} \partial_j \partial_l A_m \\
&= \left(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}\right) \partial_j \partial_l A_m \\
&= \partial_m \partial_i A_m - \partial_l \partial_l A_i = \partial_i \partial_m A_m - \partial_l \partial_l A_i \\
&= \left[\boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \mathbf{A}\right)\right]_i - (\nabla^2 \mathbf{A})_i .
\end{aligned}$$

Hence $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \mathbf{A}\right) - \nabla^2 \mathbf{A}$, since corresponding components from both sides agree.

Example 4.2.11. Show that $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\nabla \times \mathbf{B})$.

Solution 4.2.11.

$$
\begin{aligned}
\nabla \cdot (\mathbf{A} \times \mathbf{B}) &= \partial_i (\mathbf{A} \times \mathbf{B})_i = \partial_i \varepsilon_{ijk} A_j B_k = \varepsilon_{ijk} \partial_i (A_j B_k) \\
&= \varepsilon_{ijk} (\partial_i A_j) B_k + \varepsilon_{ijk} A_j (\partial_i B_k) = (\varepsilon_{ijk} \partial_i A_j) B_k - A_j (\varepsilon_{jik} \partial_i B_k) \\
&= (\nabla \times \mathbf{A})_k B_k - A_j (\nabla \times \mathbf{B})_j = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\nabla \times \mathbf{B}).
\end{aligned}
$$

Example 4.2.12. Show that

$$
\nabla \times (\mathbf{A} \times \mathbf{B}) = (\nabla \cdot \mathbf{B}) \mathbf{A} - (\nabla \cdot \mathbf{A}) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B}.
$$

Solution 4.2.12.

$$
\begin{aligned}
[\nabla \times (\mathbf{A} \times \mathbf{B})]_i &= \varepsilon_{ijk} \partial_j (\mathbf{A} \times \mathbf{B})_k = \varepsilon_{ijk} \partial_j \varepsilon_{klm} A_l B_m \\
&= \varepsilon_{kij} \varepsilon_{klm} \partial_j (A_l B_m) = \varepsilon_{kij} \varepsilon_{klm} (B_m \partial_j A_l + A_l \partial_j B_m) \\
&= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) (B_m \partial_j A_l + A_l \partial_j B_m) \\
&= B_m \partial_m A_i - B_i \partial_l A_l + A_i \partial_m B_m - A_l \partial_l B_i \\
&= [(\mathbf{B} \cdot \nabla) \mathbf{A} - (\nabla \cdot \mathbf{A}) \mathbf{B} + (\nabla \cdot \mathbf{B}) \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B}]_i .
\end{aligned}
$$

Since the corresponding components agree, the two sides of the desired equation must be equal.

4.2.7 Quotient Rule

Another way to determine if a quantity with two subscripts is a second rank tensor is to use the following quotient rule.

If for an arbitrary vector $\mathbf{B}$, the result of summing over j of the product $K_{ij} B_j$ is another vector A

$$
A_i = K_{ij} B_j, \tag{4.62}
$$

and (4.62) holds in all Cartesian coordinate systems, then K_{ij} is a true second rank tensor.

To prove the quotient rule, we examine the components of $\mathbf{A}$ in a rotated system,

$$
A_i' = a_{il} A_l = a_{il} K_{lm} B_m.
$$

Since $\mathbf{B}$ is a vector,

$$
B_m = a_{jm} B_j'.
$$

It follows

$$A'_i = a_{il}K_{lm}B_m = a_{il}K_{lm}a_{jm}B'_j = a_{il}a_{jm}K_{lm}B'_j.$$

But since (4.62) holds for all systems,

$$A'_i = K'_{ij}B'_j.$$

Subtracting the last two equations,

$$\left(K'_{ij} - a_{il}a_{jm}K_{lm}\right)B'_j = 0.$$

Since B'_j is arbitrary,

$$K'_{ij} = a_{il}a_{jm}K_{lm}.$$

Therefore K_{ij} is a second rank tensor.

With a similar procedure, one can show that if an Mth rank tensor is linearly related to an Nth rank tensor through a quantity T with $M + N$ subscripts, and the relation holds for all systems, then T is a tensor of rank $M + N$.

Example 4.2.13. If $T_{ij}x_ix_j$ is equal to a scalar S, show that T_{ij} is a second rank tensor.

Solution 4.2.13. Since x_ix_j is the outer product of two position vectors, so it is a second rank tensor. The scalar S is a zeroth rank tensor, therefore by quotient rule T_{ij} is a second rank $(2 + 0 = 2)$ tensor. It is instructive to demonstrate this directly by looking at the components in a rotated system.

$$S = T_{ij}x_ix_j = T_{lm}x_lx_m. \quad x_l = a_{il}x'_i; \quad x_m = a_{jm}x'_j.$$

$$S = T_{lm}a_{il}x'_ia_{jm}x'_j = a_{il}a_{jm}T_{lm}x'_lx'_j,$$

$$S' = T'_{ij}x'_ix'_j, \qquad S' = S.$$

Therefore

$$(T'_{ij} - a_{il}a_{jm}T_{lm})x'_ix'_j = 0; \qquad T'_{ij} = a_{il}a_{jm}T_{lm}.$$

Hence T_{ij} is a second rank tensor.

4.2.8 Symmetry Properties of Tensors

A tensor $S_{ijk\cdots}$ is said to be symmetric in the indices i and j if

$$S_{ijk\cdots} = S_{jik\cdots}.$$

A tensor $A_{ijk\cdots}$ is said to be antisymmetric in the indices i and j if

$$A_{ijk\cdots} = -A_{jik\cdots}.$$

For example, the outer product of $\mathbf{r}$ with itself $x_i x_j$ is a second rank symmetrical tensor, the Kronecker delta δ_{ij} is also a second rank symmetrical tensor. On the other hand, the Levi-Civita symbol ε_{ijk} is antisymmetric with respect to any two of its indices, since $\varepsilon_{ijk} = -\varepsilon_{jik}$.

Symmetry is a physical property of tensors. It is invariant to coordinate transformation. For example, if S_{ij} is a symmetrical tensor in certain coordinate system, in a rotated system

$$S'_{lm} = a_{li}a_{mj}S_{ij} = a_{mj}a_{li}S_{ji} = S'_{ml}.$$

Therefore S_{ij} is also a symmetric tensor in a new system. Similar results hold also for antisymmetric tensors.

A symmetric second rank tensor S_{ij} can be written in the form

$$S_{ij} = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{12} & S_{23} & S_{33} \end{pmatrix},$$

while an antisymmetric second rank tensor A_{ij} is of the form

$$A_{ij} = \begin{pmatrix} 0 & A_{12} & A_{13} \\ -A_{12} & 0 & A_{23} \\ -A_{13} & -A_{23} & 0 \end{pmatrix}.$$

Thus a symmetric second rank tensor has six independent components, while an antisymmetric second tensor has only three independent components.

Any second rank tensor T_{ij} can be represented as the sum of a symmetric tensor and an antisymmetric tensor. Given T_{ij}, one can construct

$$S_{ij} = \frac{1}{2}\left(T_{ij} + T_{ji}\right), \qquad A_{ij} = \frac{1}{2}(T_{ij} - T_{ji}).$$

Clearly S_{ij} is symmetric and A_{ij} is antisymmetric. Furthermore

$$T_{ij} = S_{ij} + A_{ij}.$$

Therefore any second rank tensor has a symmetric part and an antisymmetric part.

As shown in the theory of matrices, the six independent elements of a symmetric matrix can be represented by a quadratic surface. In the same way, a symmetric second rank tensor can be represented uniquely by an ellipsoid

$$T_{ij}x_i x_j = \pm 1,$$

where the sign is that of the determinant of $|T_{ij}|$.

The three independent components of an antisymmetric second rank tensor can also be represented geometrically by a vector.

4.2.9 Pseudotensors

One of the reasons why tensors are useful is that they make it possible to formulate the laws of physics in a way that is independent of any preferred direction in space. One might also expect these laws to be independent of whether we choose a right-handed system or a left-handed system of axes. However, under a transformation from right-handed axes to left-handed axes not all tensors transform in the same way.

So far our discussion has been restricted to rotations within right-handed systems. The right-handed system is defined by naming the three basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in such a way that the thumb of your right hand is pointing in the direction of $\mathbf{e}_3$ if the other four fingers can curl from $\mathbf{e}_1$ toward $\mathbf{e}_2$ without passing through negative $\mathbf{e}_2$. A right-handed system can be rotated into another right-handed system. The determinant of the rotation matrix is equal to 1 as shown in example 4.1.2.

Now consider the effect of inversion. The three basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are changed to $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$ such that

$$\mathbf{e}_1' = -\mathbf{e}_1, \quad \mathbf{e}_2' = -\mathbf{e}_2, \quad \mathbf{e}_3' = -\mathbf{e}_3.$$

This new set of coordinate axes is a left-handed system. This is called a left-handed system because only if you use your left hand, can the thumb point in the positive $\mathbf{e}_3'$ direction while other four fingers curl from $\mathbf{e}_1'$ toward $\mathbf{e}_2'$ without passing through negative $\mathbf{e}_2'$. Note that one cannot rotate a right-handed system into a left-handed system.

If we use the same right hand rule for the definition of vector cross products in all systems, then with right-handed axes

$$(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 = 1,$$

and with left-handed axis

$$(\mathbf{e}_1' \times \mathbf{e}_2') \cdot \mathbf{e}_3' = -1.$$

The position vector

$$\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

expressed in the inverted system becomes

$$\mathbf{r}' = -x_1\mathbf{e}_1' - x_2\mathbf{e}_2' - x_3\mathbf{e}_3'.$$

In other words, it is the same vector $\mathbf{r} = \mathbf{r}'$, except in the prime system the coefficients become negative since the axes are inverted. Vectors behaving this way when the coordinates are changed from a right-handed system to a left-handed system are called *polar vectors*. They are just regular vectors.

A fundamental difference appears when we encounter the cross product of two polar vectors. The components of $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ are given by

$$C_1 = A_2 B_3 - A_3 B_2,$$

and so on. Now when the coordinates axes are inverted, A_i goes to $-A_i$, B_i changes to $-B_i$ but C_i goes to $+C_i$ since it is the product of two negative terms. It does not behave like polar vectors under inversion. To distinguish, the cross product is called a *pseudo vector*, also known as *Axial vector*.

In addition to inversion, reflection (reversing one axis) and interchanging of two axes also transform a right-handed system into a left-handed system. The transformation matrices of these right-left operations are as follows.

Inversion:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Reflection with respect to the $x_2 x_3$ plane:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Interchange of x_1 and x_2 axes:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

These equations can be written in the form

$$x_i' = a_{ij} x_j.$$

It is obvious that the determinants $|a_{ij}|$ of these transformation matrices are all equal to -1. A left-handed system can be rotated into another left-handed system with Euler angles in the same way as in the right-handed systems. Hence, if a matrix (a_{ij}) transforms a right-handed system into a left-handed system, or vice versa, the determinant $|a_{ij}|$ is always equal to -1. Furthermore, we can show, in the same manner as with the rotation matrix, that its elements also satisfy the orthogonality condition:

$$a_{ik} a_{jk} = \delta_{ij}.$$

Therefore it is also an orthogonal transformation.

Thus the orthogonal transformations can be divided into two classes: proper transformation for which the determinant $|a_{ij}|$ is equal to one, and improper transformation for which the determinant is equal to negative one. If the transformation is a rotation, it is a proper transformation. If the transformation changes a right-handed system into a left-handed system, the transformation is improper.

A pseudovector can now be defined as satisfying the transforming rule

$$V_i' = |a_{ij}|\, a_{ij} V_j.$$

If the transformation is proper, polar vectors and pseudovectors transform in the same way. If the transformation is improper, polar vectors transform as a regular vector, but pseudovectors will flip direction.

Pseudotensors are defined in the same way. The components of a Nth rank pseudotensor transform according to the rule:

$$T'_{i_1 \cdots i_N} = |a_{ij}|\, a_{i_1 j_1} \cdots a_{i_N j_N} T_{j_1 \cdots j_N}, \tag{4.63}$$

which is exactly the same as a regular tensor, except for the determinant $|a_{ij}|$.

It follows from this definition that:

1. The outer product of two pseudotensors of rank M and N is a regular tensor of rank $M + N$.
2. The outer product of a pseudotensor of rank M and a regular tensor of rank N gives a pseudotensor of rank $M + N$.
3. The contraction of a pseudotensor of rank N gives a pseudotensor of rank $N - 2$.

A zeroth rank pseudotensor is a pseudoscalar, which changes sign under inversion, whereas a scalar does not. An example of pseudoscalar is the scalar triple product $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$. The cross product of $\mathbf{A} \times \mathbf{B}$ is a first rank pseudotensor, the polar vector $\mathbf{C}$ is a regular first rank tensor. Hence $(\mathbf{A} \times \mathbf{B})_i \cdot C_i = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$ is the contraction of a second rank pseudotensor $(\mathbf{A} \times \mathbf{B})_i \cdot C_j$, therefore a pseudoscalar. If $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are the three sides of a parallelepiped, $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$ is the volume of the parallelepiped. Defined this way, volume is actually a pseudoscalar.

We have shown ε_{ijk} is an isotropic third rank tensor under rotations. We will now show that if ε_{ijk} is to mean the same thing in both right-handed and left-handed systems, then ε_{ijk} must be regarded as third rank pseudotensor. This is because if we want $\varepsilon'_{ijk} = \varepsilon_{ijk}$ under both proper and improper transformations, then ε_{ijk} must be expressed as

$$\varepsilon'_{ijk} = |a_{ij}|\, a_{il} a_{jm} a_{nk} \varepsilon_{lmn}. \tag{4.64}$$

Since

$$a_{il} a_{jm} a_{kn} \varepsilon_{lmn} = \varepsilon_{ijk} |a_{ij}|,$$

as shown in (4.54), it follows from (4.64) that

$$\varepsilon'_{ijk} = |a_{ij}|\, \varepsilon_{ijk} |a_{ij}| = |a_{ij}|^2 \varepsilon_{ijk} = \varepsilon_{ijk}.$$

Therefore ε_{ijk} is a third rank pseudotensor.

Example 4.2.14. Let T_{12}, T_{13}, T_{23} be the three independent components of an antisymmetric tensor, show that $T_{23}, -T_{13}, T_{13}$ can be regarded as the components of a pseudovector.

Solution 4.2.14. Since ε_{ijk} is a third rank pseudotensor, T_{jk} is a second rank tensor, after contracting twice,

$$C_i = \varepsilon_{ijk} T_{jk}$$

the result C_i is a first rank pseudotensor, which is just a pseudovector. Since $T_{ij} = -T_{ji}$, so

$$T_{21} = -T_{12}, \quad T_{31} = -T_{13}, \quad T_{32} = -T_{23}.$$

Now

$$\begin{aligned} C_1 &= \varepsilon_{123} T_{23} + \varepsilon_{132} T_{32} = T_{23} - T_{32} = 2T_{23}, \\ C_2 &= \varepsilon_{213} T_{13} + \varepsilon_{231} T_{31} = -T_{13} + T_{31} = -2T_{13}, \\ C_3 &= \varepsilon_{312} T_{12} + \varepsilon_{321} T_{21} = T_{12} - T_{21} = 2T_{12}. \end{aligned}$$

Since (C_1, C_2, C_3) is a pseudovector, $(T_{23}, -T_{13}, T_{12})$ is also a pseudovector.

Example 4.2.15. Use the fact that ε_{ijk} is a third rank pseudotensor to show that $\mathbf{A} \times \mathbf{B}$ is a pseudovector.

Solution 4.2.15. Let $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, so

$$C_i = \varepsilon_{ijk} A_j B_k.$$

Expressed in a new coordinate system where $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are, respectively, transformed to $\mathbf{A}', \mathbf{B}', \mathbf{C}'$, the cross product becomes $\mathbf{C}' = \mathbf{A}' \times \mathbf{B}'$ which can be written in the component form

$$C'_l = \varepsilon'_{lmn} A'_m B'_n.$$

Since ε_{ijk} is a pseudotensor,

$$\varepsilon'_{lmn} = |a_{ij}|\, a_{li} a_{mj} a_{nk} \varepsilon_{ijk},$$

so C'_l becomes

$$C'_l = |a_{ij}|\, a_{li} a_{mj} a_{nk} \varepsilon_{ijk} A'_m B'_n.$$

But

$$A_j = a_{mj} A'_m, \qquad B_k = a_{nk} B'_n,$$

thus

$$C'_l = |a_{ij}|\, a_{li} \varepsilon_{ijk} A_j B_k = |a_{ij}|\, a_{li} C_i.$$

Therefore $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ is a pseudovector.

Since mathematical equations describing physical laws should be independent of coordinate systems, we cannot equate tensors of different rank because they have different transformation properties under rotation. Likewise, in classical physics, we cannot equate pseudotensors to tensors because they transform differently under inversion. However, surprisingly nature under the influence of weak interaction can distinguish a left-handed system from a right-handed system. By introducing a pseudoscalar, the counterintuitive events that violate parity conservation can be described. (See, T.D. Lee in "Thirty Years Since Parity Nonconservation", Birkhäuser, 1988, page 158).

4.3 Some Physical Examples

4.3.1 Moment of Inertia Tensor

One of the most familiar second rank tensors in physics is the moment of inertial tensor. It relates the angular momentum **L** and the angular velocity $\boldsymbol{\omega}$ of the rotational motion of a rigid body. The angular momentum **L** of a rigid body rotating about a fixed point is given by

$$\mathbf{L} = \int \mathbf{r} \times \mathbf{v} \mathrm{d}m,$$

where $\mathbf{r}$ is the position vector from the fixed point to the mass element $\mathrm{d}m$, and $\mathbf{v}$ is the velocity of $\mathrm{d}m$. We have shown

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r},$$

therefore

$$\begin{aligned} \mathbf{L} &= \int \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) \mathrm{d}m \\ &= \int \left[(\mathbf{r} \cdot \mathbf{r}) \boldsymbol{\omega} - (\mathbf{r} \cdot \boldsymbol{\omega}) \mathbf{r} \right] \mathrm{d}m. \end{aligned}$$

Written in tensor notation with the summation convention, the ith component of **L** is

$$L_i = \int [r^2 \omega_i - x_j \omega_j x_i] \mathrm{d}m.$$

Since

$$r^2 \omega_i = r^2 \omega_j \delta_{ij},$$

$$L_i = \omega_j \int [r^2 \delta_{ij} - x_j x_i] \mathrm{d}m = I_{ij} \omega_j,$$

where I_{ij}, known as the moment of inertia tensor, is given by

$$I_{ij} = \int [x_k x_k \delta_{ij} - x_j x_i] \mathrm{d}m. \tag{4.65}$$

Since δ_{ij} and $x_i x_j$ are both second rank tensors, I_{ij} is a symmetric second rank tensor. Explicitly the components of this tensor are

$$I_{ij} = \begin{pmatrix} \int (x_2^2 + x_3^2)\mathrm{d}m & -\int x_1 x_2 \mathrm{d}m & -\int x_1 x_3 \mathrm{d}m \\ -\int x_2 x_1 \mathrm{d}m & \int (x_1^2 + x_3^2)\mathrm{d}m & -\int x_2 x_3 \mathrm{d}m \\ -\int x_3 x_1 \mathrm{d}m & -\int x_3 x_2 \mathrm{d}m & \int (x_2^2 + x_1^2)\mathrm{d}m \end{pmatrix}. \tag{4.66}$$

4.3.2 Stress Tensor

The name tensor comes from the tensile force in elasticity theory. Inside a loaded elastic body, there are forces between neighboring parts of the material. Imagine a cut through the body, the material on the right exerts a force $\mathbf{F}$ on the material to the left, and the material on the left exerts an equal and opposite force $-\mathbf{F}$ on the material to the right.

Let us examine the force across a small area $\Delta x_1 \Delta x_3$, shown in Fig. 4.3, in the imaginary plane perpendicular to the x_2 axis. If the area is small enough, we expect the force is proportional to the area. So we can define the stress $\mathbf{P}_2$ as the force per unit area. The subscript 2 indicates that the force is acting on a plane perpendicular to the positive x_2 axis. The components of $\mathbf{P}_2$ along x_1, x_2, x_3 axes are denoted as P_{12}, P_{22}, P_{32}, respectively. Next, we can look at a small area on the imaginary plane perpendicular to the x_1 axis, and define the stress components P_{11}, P_{21}, P_{31}. Finally, imagine the cut is perpendicular to the x_3 axis, so the stress components P_{13}, P_{23}, P_{33} can be similarly defined. If $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the unit bases vectors, these relations can be expressed as

$$\mathbf{P}_j = P_{ij}\mathbf{e}_i. \tag{4.67}$$

Thus the stress has nine components

$$P_{ij} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}. \tag{4.68}$$

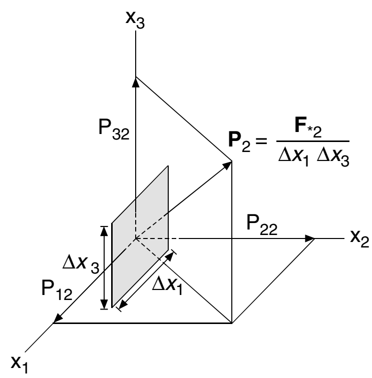

Fig. 4.3. Stress $\mathbf{P}_2$, defined as force per unit area, on a small surface perpendicular to the x_2 axis. Its components along the three axes are, respectively, P_{12}, P_{22}, P_{32}

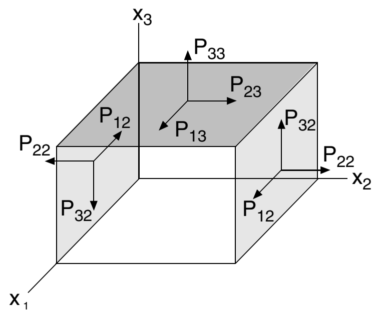

Fig. 4.4. The nine components of a stress tensor at a point can be represented as normal and tangential forces on the surfaces of an infinitesimal cube around the point

The first subscript in P_{ij} indicates the direction of the force component, the second subscript indicates the direction of the normal to the surface on which the force is acted upon.

The physical meaning of P_{ij} is as follows. Imagine an infinitesimal cube around a point inside the material, P_{ij} are the forces per unit area on the faces of this cube, as shown in Fig. 4.4. For clarity, forces are shown only on three surfaces. There are both normal forces P_{ii} (tensions shown but they could be pressures with arrows reversed) and tangential forces P_{ij} ($i \neq j$, shears). Note that, in equilibrium, the forces on the opposite faces must be equal and opposite. Furthermore, (4.67) is symmetric $P_{ij} = P_{ji}$, because of rotational equilibrium. For example, the shearing force on the top surface in the x_2 direction is $P_{23}\Delta x_1 \Delta x_2$. The torque around x_1 axis due to this force is $(P_{23}\Delta x_1 \Delta x_2)\Delta x_3$. The opposite torque due to the shearing force on the right surface is $(P_{32}\Delta x_3 \Delta x_1)\Delta x_2$. Since the net torque around x_1 axis must be zero, therefore

$$(P_{23}\Delta x_1 \Delta x_2)\Delta x_3 = (P_{32}\Delta x_3 \Delta x_1)\Delta x_1,$$

and we have

$$P_{23} = P_{32}.$$

Similar argument will show that in general $P_{ij} = P_{ji}$, therefore (4.68) is symmetric.

Now we are going to show that the nine components of (4.68) is a tensor, known as the *stress tensor.* For this purpose, we construct an infinitesimal tetrahedron with its edges directed along the coordinate axes as shown in Fig. 4.5. Let $\Delta a_1, \Delta a_2, \Delta a_3$ denote the areas of the faces perpendicular to the axes x_1, x_2, x_3, respectively, the forces per unit area on these faces are $-\mathbf{P}_1, -\mathbf{P}_2, -\mathbf{P}_3$, since these faces are directed in the negative direction of the axes. Let Δa_n denote the area of the inclined face with an unit exterior normal $\mathbf{n}$, and $\mathbf{P}_n$ be the force per unit area on this surface. The total force on these

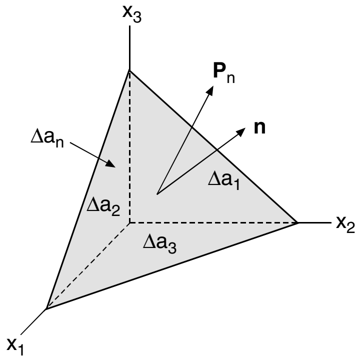

Fig. 4.5. Forces on the surfaces of an infinitesimal tetrahedron. The equilibrium conditions show that the stress must be a second rank tensor

four surfaces must be zero, even if there are body forces, such as gravity. The body forces will be proportional to the volume, whereas all the surface forces are proportional to the area. As dimensions shrink to zero, the body forces have one more infinitesimal, and can always be neglected compared with the surface forces. Therefore

$$\mathbf{P}_n \Delta a_n - \mathbf{P}_1 \Delta a_1 - \mathbf{P}_2 \Delta a_2 - \mathbf{P}_3 \Delta a_3 = \mathbf{0}. \tag{4.69}$$

Since Δa_1 is the area of Δa_n projected on the $x_2 x_3$ plane, therefore

$$\Delta a_1 = (\mathbf{n} \cdot \mathbf{e}_1)\, \Delta a_n.$$

With similar expressions of Δa_2 and Δa_3, we can write (4.69) in the form

$$\mathbf{P}_n \Delta a_n = \mathbf{P}_1 (\mathbf{n} \cdot \mathbf{e}_1)\, \Delta a_n + \mathbf{P}_2 (\mathbf{n} \cdot \mathbf{e}_2)\, \Delta a_n + \mathbf{P}_3 (\mathbf{n} \cdot \mathbf{e}_3)\, \Delta a_n,$$

or

$$\mathbf{P}_n = (\mathbf{n} \cdot \mathbf{e}_k)\, \mathbf{P}_k. \tag{4.70}$$

What we want to find is the stress components in a rotated system with axes $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$. Now without loss of generality, we can assume that the jth axis of the rotated system is directed along $\mathbf{n}$, that is

$$\mathbf{n} = \mathbf{e}'_j.$$

Therefore $\mathbf{P}_n$ is $\mathbf{P}'_j$ in the rotated system. Thus (4.70) becomes

$$\mathbf{P}'_j = \left(\mathbf{e}'_j \cdot \mathbf{e}_k\right) \mathbf{P}_k. \tag{4.71}$$

In terms of their components along the respective coordinate axes, as shown in (4.67),

$$\mathbf{P}'_j = P'_{ij} \mathbf{e}'_i, \qquad \mathbf{P}_k = P_{lk} \mathbf{e}_l$$

the last equation can be written as

$$P'_{ij}\mathbf{e}'_i = \left(\mathbf{e}'_j \cdot \mathbf{e}_k\right) P_{lk}\mathbf{e}_l.$$

Take the dot product with $\mathbf{e}'_i$ on both sides,

$$P'_{ij}(\mathbf{e}'_i \cdot \mathbf{e}'_i) = \left(\mathbf{e}'_j \cdot \mathbf{e}_k\right) P_{lk}(\mathbf{e}_l \cdot \mathbf{e}'_i),$$

we have

$$P'_{ij} = (\mathbf{e}'_i \cdot \mathbf{e}_l)\left(\mathbf{e}'_j \cdot \mathbf{e}_k\right) P_{lk}.$$

Since $(\mathbf{e}'_m \cdot \mathbf{e}_n) = a_{mn}$, we see that

$$P'_{ij} = a_{il}a_{jk}P_{lk}. \tag{4.72}$$

Therefore the array of the stress components (4.68) is indeed a tensor.

4.3.3 Strain Tensor and Hooke's Law

Under imposed forces, an elastic body will deform and exhibit strain. The deformation is characterized by the change of distances between neighboring points. Let P at $\mathbf{r}$ and Q at $\mathbf{r} + \boldsymbol{\Delta}\mathbf{r}$ be two nearby points, as shown in Fig. 4.6. When the body is deformed, point P is displaced by the amount $\mathbf{u}(\mathbf{r})$ to P', and Q by $\mathbf{u}(\mathbf{r} + \boldsymbol{\Delta}\mathbf{r})$ to Q'. If the displacements of these neighboring points are the same, that is if $\mathbf{u}(\mathbf{r}) = \mathbf{u}(\mathbf{r} + \boldsymbol{\Delta}\mathbf{r})$, the relative positions of the points are not changed. That part of the body is unstrained, since the distances PQ and $P'Q'$ will be the same. Therefore the strain is associated with the variation of the displacement vector $\mathbf{u}(\mathbf{r})$. The change of $\mathbf{u}(\mathbf{r})$ can be written as

$$\boldsymbol{\Delta}\mathbf{u} = \mathbf{u}(\mathbf{r} + \boldsymbol{\Delta}\mathbf{r}) - \mathbf{u}(\mathbf{r}).$$

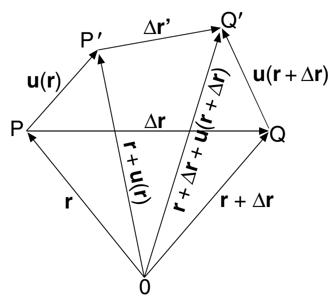

Fig. 4.6. The strain of a deformed elastic body. The body is strained if the relative distance of two nearby points is changed. The strain tensor depends on the variation of the displacement vector $\mathbf{u}$ with respect to the position vector $\mathbf{r}$

Neglecting second and higher terms, the components of $\mathbf{\Delta u}$ can be written as

$$\begin{aligned}\Delta u_i &= u_i\left(x_1+\Delta x_1, x_2+\Delta x_2, x_3+\Delta x_3\right) - u_i\left(x_1, x_2, x_3\right) \\ &= \frac{\partial u_i}{\partial x_1}\Delta x_1 + \frac{\partial u_i}{\partial x_2}\Delta x_2 + \frac{\partial u_i}{\partial x_3}\Delta x_3 = \frac{\partial u_i}{\partial x_j}\Delta x_j.\end{aligned}$$

Since u_i is a vector and $\partial/\partial x_j$ is a vector operator, $\partial u_i/\partial x_j$ is the outer product of two first rank tensors. Therefore $\partial u_i/\partial x_j$ is a second rank tensor. It can be decomposed into a symmetric part and an antisymmetric part

$$\frac{\partial u_i}{\partial x_j} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right). \tag{4.73}$$

We can also divide $\mathbf{\Delta u}$ into two parts

$$\mathbf{\Delta u} = \mathbf{\Delta u}^{\mathrm{s}} + \mathbf{\Delta u}^{\mathrm{a}},$$

where

$$\Delta u_i^{\mathrm{s}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\Delta x_j,$$

$$\Delta u_i^{\mathrm{a}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)\Delta x_j.$$

The antisymmetric part of (4.73) does not alter the distance between P and Q because of the following. Let the distance $P'Q'$ be $\mathbf{\Delta r}'$. It is clear from Fig. 4.6 that

$$\mathbf{\Delta r}' = \left[\mathbf{\Delta r} + \mathbf{u}\left(\mathbf{r}+\mathbf{\Delta r}\right)\right] - \mathbf{u}\left(\mathbf{r}\right).$$

It follows that

$$\mathbf{\Delta r}' - \mathbf{\Delta r} = \mathbf{u}\left(\mathbf{r}+\mathbf{\Delta r}\right) - \mathbf{u}\left(\mathbf{r}\right) = \mathbf{\Delta u} = \mathbf{\Delta u}^{s} + \mathbf{\Delta u}^{a}.$$

Now

$$\mathbf{\Delta u}^{a}\cdot\mathbf{\Delta r} = \Delta u_i^{a}\Delta x_i = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)\Delta x_j \Delta x_i = 0.$$

This is because in this expression, both i and j are summing indices and can be interchanged. Thus $\mathbf{\Delta u}^{\mathrm{a}}$ is perpendicular to $\mathbf{\Delta r}$, and it can be considered as the infinitesimal arc length of a rotation around the tail of $\mathbf{\Delta r}$, as shown in the following sketch.

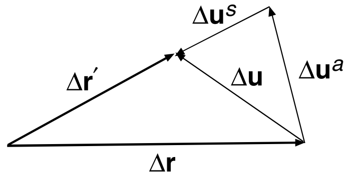

Fig. 4.7. The change of distance between two nearby points in an elastic body. It is determined by the symmetric strain tensor

Therefore $\mathbf{\Delta u}^{\mathrm{a}}$ does not change the length $\Delta\mathbf{r}$. The change of distance between two nearby points of an elastic body is uniquely determined by the symmetric part of (4.73)

$$E_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{4.74}$$

This quantity is known as the strain tensor. The stain tensor plays an important role in the elasticity theory because it is a measure of the degree of deformation.

Since in one-dimension the elastic force in a spring is given by the Hooke's law $F = -kx$, one might expect that in three-dimensional elastic media the strain is proportional to the stress. For most solid materials with a relative strain of a few percent, this is indeed the case. The linear relationship between the strain tensor and the stress tensor is given by the generalized Hooke's law

$$P_{ij} = c_{ijkl}E_{kl}, \tag{4.75}$$

where c_{ijkl} is known as the elasticity tensor. Since P_{ij} and E_{kl} are both second rank tensors, by the quotient rule, c_{ijkl} must be a fourth rank tensor. While there are 81 components of a fourth rank tensor, but various symmetry considerations will show that the number of independent components in a general crystalline body is only 21. If the body is isotropic, the elastic constants are further reduced to only two. While we are not going into these details which are the subjects of books on elasticity, we only want to show that concepts of tensor are crucial in describing these physical quantities.

Exercises

1. Find the rotation matrix for
   (a) a rotation of $\pi/2$ about z axis,
   (b) a rotation of π about x axis.
   Ans. $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

2. With the transformation matrix (A) given by (4.20), show that
   $$(A)\left(A^{\mathrm{T}}\right) = (I),$$
   where (I) is the identity matrix.

3. With the transformation matrix (A) given by (4.20), explicitly verify that
   $$\sum_{i=1}^{3} a_{ij}a_{ik} = \delta_{jk},$$
   for (a) $j = 1,\ k = 1$, (b) $j = 1,\ k = 2$, (c) $j = 1,\ k = 3$.

4. With the transformation matrix (A) given by (4.20), show explicitly that the determinant A is equal to 1.

5. Show that there is no nontrivial isotropic first rank tensor.
Hint: (1) Assume there is an isotropic first rank tensor (A_1, A_2, A_3). Under a rotation, $A_1' = A_1,\ A_2' = A_2,\ A_3' = A_3$, since it is isotropic. (2) Consider a rotation of $\pi/2$ about x_3 axis, and show that $A_1 = 0,\ A_2 = 0$. (3) A rotation about x_1 axis will show that $A_3 = 0$. (4) Therefore only the zero vector is a first rank isotropic tensor.

6. Let
$$T_{ij} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix}, \quad A_i = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$
Find the following contractions
(a) $B_i = T_{ij} A_j$,
(b) $C_j = T_{ij} A_i$,
(c) $S = T_{ij} A_i A_j$.
Ans. (a) $B_i = (5, 5, 10)$, (b) $C_i = (4, 6, 11)$, (c) $S = 35$.

7. Let A_{ij} and B_{ij} be two second rank tensors, and let
$$C_{ij} = A_{ij} + B_{ij}.$$
Show that C_{ij} is also a second rank tensor.

8. The equation of an ellipsoid centered at the origin is of the form
$$A_{ij} x_i x_j = 1.$$
Show that A_{ij} is a second rank tensor.
Hint: In a rotated system, the equation of the surface is $A'_{ij} x'_i x'_j = 1$.

9. Show that
$$A_i = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$
is a two-dimensional vector.

10. Show that the following 2×2 matrices represent second rank tensors in two-dimensional space:
$$\text{(a)} \begin{pmatrix} -x_1 x_2 & x_1^2 \\ -x_2^2 & x_1 x_2 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} x_2^2 & -x_1 x_2 \\ -x_1 x_2 & x_1^2 \end{pmatrix},$$
$$\text{(c)} \begin{pmatrix} -x_1 x_2 & -x_2^2 \\ x_1^2 & x_1 x_2 \end{pmatrix}, \quad \text{(d)} \begin{pmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{pmatrix}.$$
Hint: Show that they are various outer products of the position vector and the vector in the last question.

11. Explicitly show that

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn} = \begin{vmatrix} a_{i1} & a_{i2} & a_{i3} \\ a_{j1} & a_{j2} & a_{j3} \\ a_{k1} & a_{k2} & a_{k3} \end{vmatrix}$$

by (a) writing out all nonzero terms of

$$\sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}a_{jm}a_{kn}\varepsilon_{lmn},$$

and (b) expand the determinant

$$\begin{vmatrix} a_{i1} & a_{i2} & a_{i3} \\ a_{j1} & a_{j2} & a_{j3} \\ a_{k1} & a_{k2} & a_{k3} \end{vmatrix}$$

over the elements of the first row.

12. Show that ε_{ijk} can be written as

$$\varepsilon_{ijk} = \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{vmatrix}.$$

Hint: Show that the determinant has all the properties ε_{ijk} has.

13. Show that
(a) $\sum_{ij} \varepsilon_{ijk}\delta_{ij} = 0$;
(b) $\sum_{jk} \varepsilon_{ijk}\varepsilon_{ljk} = 2\delta_{il}$;
(c) $\sum_{ijk} \varepsilon_{ijk}\varepsilon_{ijk} = 6$.

14. The following equations are written with summation convention, verify them
(a) $\delta_{ij}\delta_{jk}\delta_{ki} = 3$,
(b) $\varepsilon_{ijk}\varepsilon_{klm}\varepsilon_{mni} = \varepsilon_{jnl}$.
Hint: (a) Recall δ_{ij} is a substitution tensor, (b) Use (4.60).

15. With the summation convention, show that
(a) $A_i\delta_{ij} = A_j$
(b) $B_j\delta_{ij} = B_i$
(c) $\delta_{1j}\delta_{j1} = 1$
(d) $\delta_{ij}\delta_{ji} = \delta_{ii} = 3$
(e) $\delta_{ij}\delta_{jl} = \delta_{il}$

16. With subscripts and summation convention, show that
(a) $\partial_i x_j = \delta_{ij}$
(b) $\partial_i \left(x_j x_j\right)^{1/2} = \dfrac{1}{\left(x_j x_j\right)^{1/2}} x_i$.
Hint: (a) x_1, x_2, x_3 are independent variables. (b) $\partial_i \left(x_j x_j\right) = 2x_j \partial_i x_j$.

17. The following equations are written with summation convention (a) $\partial_i x_i = 3$, (b) $\partial_i (x_j x_j)^{1/2} = (x_j x_j)^{-1/2} x_i$, translate them into ordinary vector notation.
Ans. (a) $\boldsymbol{\nabla} \cdot \mathbf{r} = 3$, (b) $\boldsymbol{\nabla} r = \mathbf{r}/r$.

18. The following expressions are written with summation convention
(a) $V_i A_j B_i \mathbf{e}_j$; (b) $c A_i B_j \delta_{ij}$;
(c) $A_l B_j \varepsilon_{ijk} \delta_{li} \mathbf{e}_k$; (d) $\varepsilon_{ijk} \varepsilon_{lmk} A_i B_j C_l D_m$,
translate them into ordinary vector notation.
Ans. (a) $(\mathbf{V} \cdot \mathbf{B})\, \mathbf{A}$, (b) $c\mathbf{A} \cdot \mathbf{B}$, (c) $\mathbf{A} \times \mathbf{B}$, (d) $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D})$.

19. Use the Levi-Civita tensor technique to prove the following identities:
(a) $\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}$,
(b) $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$.

20. Use the Levi-Civita tensor technique to prove the following identity

$$\boldsymbol{\nabla} \times (\phi \mathbf{A}) = \phi\, (\boldsymbol{\nabla} \times \mathbf{A}) + (\boldsymbol{\nabla} \phi) \times \mathbf{A}.$$

21. Let

$$T_{ij} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}.$$

Find the symmetric part S_{ij} and the antisymmetric part A_{ij} of the tensor T_{ij}.

Ans. $S_{ij} = \begin{pmatrix} 1 & 1 & 1.5 \\ 1 & 4 & 2.5 \\ 1.5 & 2.5 & 6 \end{pmatrix}, \quad A_{ij} = \begin{pmatrix} 0 & 1 & 1.5 \\ -1 & 0 & 2.5 \\ -1.5 & -2.5 & 0 \end{pmatrix}.$

22. If S_{ij} is a symmetric tensor and A_{ij} is an antisymmetric tensor, show that

$$S_{ij} A_{ij} = 0.$$

23. Let φ be a scalar, V_i be a pseudovector, T_{ij} be a second rank tensor, and let

$$A_{ijk} = \varepsilon_{ijk} \varphi, \quad B_{ij} = \varepsilon_{ijk} V_k, \quad C_i = \varepsilon_{ijk} T_{jk}.$$

Show that A_{ijk} is a third rank pseudotensor, B_{ij} is a second rank tensor, and C_i is a pseudovector.

24. Find the strain tensor for an isotropic elastic material when it is subjected to
(a) A stretching deformation $\mathbf{u} = (0, 0, \alpha x_3)$;
(b) A shearing deformation $\mathbf{u} = (\beta x_3, 0, 0)$.

Ans. (a) $E_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$; (b) $E_{ij} = \begin{pmatrix} 0 & 0 & \beta/2 \\ 0 & 0 & 0 \\ \beta/2 & 0 & 0 \end{pmatrix}$.