1
Suffix notation and tensor algebra

1.1 Summation convention

We consider a rectangular cartesian coordinate system with unit vectors **i**, **j** and **k** along the three coordinate axes x, y and z respectively. For convenience, we relabel these unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ and denote the coordinate axes by x_1, x_2 and x_3 (Figure 1.1). A typical vector **a** with components a_1, a_2 and a_3 in cartesian coordinates can then be written as

$$\mathbf{a} = \sum_{i=1}^{3} a_i \mathbf{e}_i. \tag{1.1}$$

Instead of writing the summation sign in (1.1) every time we have an expression of this kind, we can adopt the summation convention: whenever an index occurs precisely twice in a term, it is understood that the index is to be summed over its full range of possible values without the need for explicitly writing the summation sign $\sum$. Hence (1.1), with this convention, is

$$\mathbf{a} = a_i \mathbf{e}_i, \tag{1.2}$$

where summation over i is implied ($i = 1, 2, 3$). Since the components a_i of **a** are given by the dot-product of **a** with each of the unit vectors $\mathbf{e}_i$, then $a_i = \mathbf{a} \cdot \mathbf{e}_i$ ($i = 1, 2, 3$) and (1.2) can be written

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{e}_i)\mathbf{e}_i, \tag{1.3}$$

again adopting the summation convention.

Example 1 If vectors **c** and **d** have components c_i and d_i $(i = 1, 2, 3)$ respectively, then

$$\mathbf{c}\,.\,\mathbf{d} = \sum_{i=1}^{3} c_i d_i = c_i d_i. \quad \blacktriangle \tag{1.4}$$

Example 2

$$\sum_{s=1}^{2} a_s x_s = a_1 x_1 + a_2 x_2 = a_i x_i. \quad \blacktriangle \tag{1.5}$$

Example 3 Consider the term $a_{ij} b_i b_j$ in which i and j both occur twice. The summation convention implies summation over i and j independently. Hence, if i and j run from 1 to 2,

$$a_{ij} b_i b_j = a_{11} b_1 b_1 + a_{12} b_1 b_2 + a_{21} b_2 b_1 + a_{22} b_2 b_2 \tag{1.6}$$

$$= a_{11} b_1^2 + a_{22} b_2^2 + b_1 b_2 (a_{12} + a_{21}). \quad \blacktriangle \tag{1.7}$$

The usual rules apply when the summation convention is being used, that is,

$$a_i b_i = b_i a_i \tag{1.8}$$

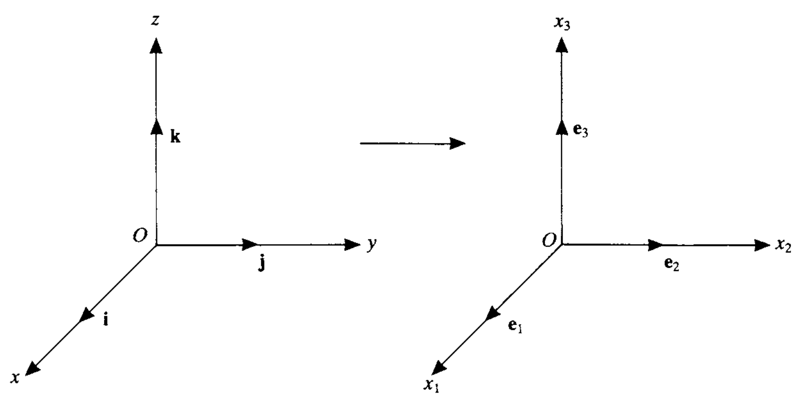
and

$$a_i (b_i + c_i) = a_i b_i + a_i c_i. \tag{1.9}$$

We see in (1.9) that although i occurs three times in the left-hand side it only occurs twice in each *term* and therefore summation over i is implied. Also

$$(a_i b_i)(c_j d_j) = (a_i c_j)(b_i d_j), \tag{1.10}$$

Figure 1.1



that is, the order of the factors is unimportant and summation is implied over both i and j. It would be wrong, however, to write (1.10) as $a_i b_i c_i d_i$ since the index occurs four times and therefore no summation over i would be implied.

Example 4

$$\left(\sum_{s=1}^{2} a_s x_s\right)^2 = (a_1 x_1 + a_2 x_2)^2 \tag{1.11}$$

$$= a_1^2 x_1^2 + 2a_1 a_2 x_1 x_2 + a_2^2 x_2^2. \tag{1.12}$$

Using the summation convention we can write this as

$$\left(\sum_{i=1}^{2} a_i x_i\right)\left(\sum_{j=1}^{2} a_j x_j\right) = a_i x_i a_j x_j, \tag{1.13}$$

where both i and j appear twice and consequently are both summed over the values 1 and 2. ◢

1.2 Free and dummy indices

Consider the following set of n linear equations for the quantities $x_1, x_2, \ldots, x_n$ with (constant) coefficients $a_{11}, a_{12}, \ldots, a_{1n}$, $a_{2n}, \ldots, a_{nn}$:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= c_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= c_2, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= c_n, \end{aligned}\right\} \tag{1.14}$$

where $c_1, c_2, \ldots, c_n$ are given constants. This set of equations can be written as

$$\left.\begin{aligned} \sum_{j=1}^{n} a_{1j}x_j &= c_1, \\ \sum_{j=1}^{n} a_{2j}x_j &= c_2, \\ \vdots \qquad & \quad \vdots \\ \sum_{j=1}^{n} a_{nj}x_j &= c_n. \end{aligned}\right\} \tag{1.15}$$

By introducing the index i, (1.15) can be expressed in the more compact form

$$\sum_{j=1}^{n} a_{ij}x_j = c_i, \quad (i = 1, 2, \ldots, n). \tag{1.16}$$

Using the summation convention, we can write (1.16) as

$$a_{ij}x_j = c_i. \tag{1.17}$$

Indices which occur twice, so that summation over them is implied (j in (1.17)), are called dummy indices, while indices which can have any value (i in (1.17)) are called free indices. We note that (1.17) could equally have been written as $a_{pq}x_q = c_p$, where summation over the dummy index q is implied and p is free.

Example 5 Suppose we are given constants a_{ij} ($i, j = 1, 2, 3$) and the function $\phi = a_{ij}x_ix_j$ (where summation over both i and j is implied). We wish to calculate the quantities $\partial\phi/\partial x_s$, where s is a free index (equal to 1, 2 or 3). Then by the chain rule for differentiation

$$\frac{\partial\phi}{\partial x_s} = a_{ij}\frac{\partial x_i}{\partial x_s}x_j + a_{ij}x_i\frac{\partial x_j}{\partial x_s}. \tag{1.18}$$

Since the x_i are independent variables, $\partial x_i/\partial x_s$ is 1 if $i = s$ and zero otherwise. Hence in the first term the i-summation has only one non-zero term (when $i = s$). Similarly the j-summation in the second term has only one non-zero term (when $j = s$). Thus

$$\partial\phi/\partial x_s = a_{sj}x_j + a_{is}x_i, \tag{1.19}$$

where j is summed in the first term and i in the second. These dummy indices can be given any letter we choose so that, replacing the dummy index j by i in the first term of (1.19),

$$\frac{\partial\phi}{\partial x_s} = a_{si}x_i + a_{is}x_i = (a_{si} + a_{is})x_i, \tag{1.20}$$

where i is now the dummy index and s is the free index. If a_{is} is symmetric so that $a_{is} = a_{si}$ then

$$\partial\phi/\partial x_s = 2a_{is}x_i, \tag{1.21}$$

whereas if a_{is} is skew- (or anti-) symmetric so that $a_{is} = -a_{si}$ then

$$\phi = 0 \quad \text{and} \quad \partial\phi/\partial x_s = 0. \tag{1.22}$$

1.3 Special symbols

1. Kronecker delta

The Kronecker delta symbol δ_{ij} is defined by

$$\delta_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases} \tag{1.23}$$

We note that $\delta_{ij}\mathbf{e}_j = \delta_{i1}\mathbf{e}_1 + \delta_{i2}\mathbf{e}_2 + \delta_{i3}\mathbf{e}_3$ and that only one of these three terms is non-zero depending on i; for example $\delta_{11} = 1$, $\delta_{12} = \delta_{13} = 0$. Hence $\delta_{ij}\mathbf{e}_j = \mathbf{e}_i$. Also, as in (1.3),

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{e}_i)\mathbf{e}_i = (\mathbf{a} \cdot \mathbf{e}_j\delta_{ij})(\mathbf{e}_k\delta_{ik}). \tag{1.24}$$

Consider the quantity $\delta_{ij}\delta_{jk}$. Then

$$\delta_{ij}\delta_{jk} = (\mathbf{e}_i \cdot \mathbf{e}_j)(\mathbf{e}_j \cdot \mathbf{e}_k) \tag{1.25}$$

$$= (\mathbf{e}_i \cdot \mathbf{e}_1)(\mathbf{e}_1 \cdot \mathbf{e}_k) + (\mathbf{e}_i \cdot \mathbf{e}_2)(\mathbf{e}_2 \cdot \mathbf{e}_k) + (\mathbf{e}_i \cdot \mathbf{e}_3)(\mathbf{e}_3 \cdot \mathbf{e}_k). \tag{1.26}$$

Now if $i \neq k$, then at most one of the two brackets in each term can be non-zero and hence each term is zero. If $i = k$ then one term has both brackets non-zero and equal to 1. Hence $\delta_{ij}\delta_{jk}$ is zero if $i \neq k$ and is 1 if $i = k$. This is just the definition of δ_{ik} (see (1.23)) and so

$$\delta_{ij}\delta_{jk} = \delta_{ik}. \tag{1.27}$$

Further consider the expression $\delta_{rs}A_{pqs}$, where r, p and q are free indices. Then

$$\delta_{rs}A_{pqs} = \delta_{r1}A_{pq1} + \delta_{r2}A_{pq2} + \delta_{r3}A_{pq3}. \tag{1.28}$$

Only one of these terms is non-zero depending on the value of r. Hence

$$\delta_{rs}A_{pqs} = A_{pqr}. \tag{1.29}$$

2. The alternating symbol

The alternating symbol ϵ_{ijk} is defined as

$$\epsilon_{ijk} = \mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k). \tag{1.30}$$

Hence if i, j and k are all different, then $\mathbf{e}_j \times \mathbf{e}_k = \pm\mathbf{e}_i$, the plus sign being taken if i, j, k form a cyclic permutation of 1, 2, 3 (1, 2, 3 or 3, 1, 2 or 2, 3, 1) and the minus sign if they form an anticyclic permutation (3, 2, 1 or 2, 1, 3 or 1, 3, 2). Hence, from (1.30), $\epsilon_{ijk} = +1$ if i, j, k are all different and cyclic, and $\epsilon_{ijk} = -1$ if i, j, k are all different and anticyclic. If $j = k$, then the cross-product is zero in (1.30) and consequently so is ϵ_{ijk}. If either j or k equals i then $\mathbf{e}_j \times \mathbf{e}_k$ is at right-angles to $\mathbf{e}_i$ and ϵ_{ijk} will again be zero. We have finally

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } i, j, k \text{ are a cyclic permutation of } 1, 2, 3, \\ -1 & \text{if } i, j, k \text{ are an anticyclic permutation of } 1, 2, 3, \\ 0 & \text{if any two (or all) indices are equal.} \end{cases} \tag{1.31}$$

Hence cyclically permuting the indices i, j, k leaves ϵ_{ijk} unaffected, whereas interchanging any two indices changes its sign:

$$\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki}, \tag{1.32}$$

$$\epsilon_{jik} = -\epsilon_{ijk}. \tag{1.33}$$

1.4 Vector identities

Consider the vector product $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Then from (1.2)

$$\mathbf{c} = (a_i\mathbf{e}_i) \times (b_j\mathbf{e}_j) = a_ib_j(\mathbf{e}_i \times \mathbf{e}_j). \tag{1.34}$$

Now, in general, $\mathbf{c} = (\mathbf{c} \,.\, \mathbf{e}_k)\mathbf{e}_k$ using (1.3). Hence substituting for $\mathbf{c}$ from (1.34) gives

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = (a_ib_j(\mathbf{e}_i \times \mathbf{e}_j) \,.\, \mathbf{e}_k)\mathbf{e}_k \tag{1.35}$$

$$= a_ib_j\epsilon_{kij}\mathbf{e}_k, \tag{1.36}$$

using (1.30). In (1.36), summation over the indices i, j and k is implied so that

$$\mathbf{a} \times \mathbf{b} = a_ib_j\epsilon_{1ij}\mathbf{e}_1 + a_ib_j\epsilon_{2ij}\mathbf{e}_2 + a_ib_j\epsilon_{3ij}\mathbf{e}_3 \tag{1.37}$$

and the rth component of $\mathbf{a} \times \mathbf{b}$ is therefore

$$(\mathbf{a} \times \mathbf{b})_r = \epsilon_{rij}a_ib_j, \quad (r = 1, 2, 3). \tag{1.38}$$

For the scalar triple product

$$(\mathbf{a} \times \mathbf{b}) \,.\, \mathbf{c} = (a_ib_j\mathbf{e}_i \times \mathbf{e}_j) \,.\, c_k\mathbf{e}_k \tag{1.39}$$

$$= a_ib_jc_k\mathbf{e}_k \,.\, (\mathbf{e}_i \times \mathbf{e}_j) \tag{1.40}$$

$$= \epsilon_{kij}a_ib_jc_k. \tag{1.41}$$

Using (1.32), we have finally

$$(\mathbf{a} \times \mathbf{b}) \,.\, \mathbf{c} = \epsilon_{ijk}a_ib_jc_k. \tag{1.42}$$

The scalar triple product does not depend on the order of the dot and cross operations since $\mathbf{a} \,.\, (\mathbf{b} \times \mathbf{c}) = (\mathbf{b} \times \mathbf{c}) \,.\, \mathbf{a} = \epsilon_{pqr}b_pc_qa_r = \epsilon_{rpq}a_rb_pc_q = (\mathbf{a} \times \mathbf{b}) \,.\, \mathbf{c}$.

We can use a result from vector algebra to derive an important identity involving the Kronecker delta and the alternating symbol. We have the standard result

$$(\mathbf{a} \times \mathbf{b}) \,.\, (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \,.\, \mathbf{c})(\mathbf{b} \,.\, \mathbf{d}) - (\mathbf{a} \,.\, \mathbf{d})(\mathbf{b} \,.\, \mathbf{c}). \tag{1.43}$$

Putting $\mathbf{a} = \mathbf{e}_i$, $\mathbf{b} = \mathbf{e}_j$, $\mathbf{c} = \mathbf{e}_k$ and $\mathbf{d} = \mathbf{e}_l$, then

$$(\mathbf{e}_i \times \mathbf{e}_j) \,.\, (\mathbf{e}_k \times \mathbf{e}_l) = (\mathbf{e}_i \,.\, \mathbf{e}_k)(\mathbf{e}_j \,.\, \mathbf{e}_l) - (\mathbf{e}_i \,.\, \mathbf{e}_l)(\mathbf{e}_j \,.\, \mathbf{e}_k). \tag{1.44}$$

Now since for any vector $\mathbf{A}$, $\mathbf{A} = (\mathbf{A} \cdot \mathbf{e}_m)\mathbf{e}_m$ (as in (1.3)),

$$(\mathbf{e}_i \times \mathbf{e}_j) = [(\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_m]\mathbf{e}_m = \epsilon_{mij}\mathbf{e}_m = \epsilon_{ijm}\mathbf{e}_m. \tag{1.45}$$

Similarly

$$(\mathbf{e}_k \times \mathbf{e}_l) = \epsilon_{klp}\mathbf{e}_p. \tag{1.46}$$

Hence the left-hand side of (1.44) becomes

$$(\mathbf{e}_i \times \mathbf{e}_j) \cdot (\mathbf{e}_k \times \mathbf{e}_l) = \epsilon_{ijm}\epsilon_{klp}\mathbf{e}_m \cdot \mathbf{e}_p \tag{1.47}$$

$$= \epsilon_{ijm}\epsilon_{klp}\delta_{mp} \tag{1.48}$$

$$= \epsilon_{ijm}\epsilon_{klm}, \tag{1.49}$$

using (1.23) and (1.29). Substituting this into (1.44) and expressing all the dot-products on the right-hand side using (1.23), we have

$$\epsilon_{ijm}\epsilon_{klm} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}. \tag{1.50}$$

Example 6 A matrix ϕ has elements

$$\phi_{ik} = n_i n_k + \epsilon_{ilk} n_l, \tag{1.51}$$

where n_i are the components of a unit vector. Show that the elements of the matrix ϕ^2 are given by

$$(\phi^2)_{ij} = 2n_i n_j - \delta_{ij}. \tag{1.52}$$

Now

$$(\phi^2)_{ij} = \phi_{ik}\phi_{kj} = (n_i n_k + \epsilon_{ilk} n_l)(n_k n_j + \epsilon_{kmj} n_m) \tag{1.53}$$

$$= n_i n_j n_k n_k + (\epsilon_{ilk} n_l n_k) n_j$$
$$+ (\epsilon_{kmj} n_k n_m) n_i + \epsilon_{ilk}\epsilon_{kmj} n_l n_m. \tag{1.54}$$

Since ϵ_{pqr} is an antisymmetric symbol under interchange of any two indices, quantities such as $\epsilon_{ilk} n_l n_k$ are zero because for any pair of values l and k, two terms (with opposite signs) result from the summations over l and k (for example, $\epsilon_{i12} n_1 n_2$ cancels with $\epsilon_{i21} n_2 n_1 = -\epsilon_{i12} n_1 n_2$). Hence, since $n_k n_k = 1$,

$$(\phi^2)_{ij} = n_i n_j + \epsilon_{ilk}\epsilon_{kmj} n_l n_m \tag{1.55}$$

$$= n_i n_j + \epsilon_{ilk}\epsilon_{mjk} n_l n_m. \tag{1.56}$$

Using (1.50), (1.56) becomes

$$(\phi^2)_{ij} = n_i n_j + (\delta_{im}\delta_{lj} - \delta_{ij}\delta_{lm}) n_l n_m \tag{1.57}$$

$$= n_i n_j + n_j n_i - \delta_{ij} n_l n_l = 2n_i n_j - \delta_{ij}, \tag{1.58}$$

as required. ◢

1.5 Vector operators

Defining the operator ∇_i by

$$\nabla_i = \frac{\partial}{\partial x_i} \tag{1.59}$$

and the vector operator $\boldsymbol{\nabla}$ by

$$\boldsymbol{\nabla} = \mathbf{e}_i \nabla_i, \tag{1.60}$$

the gradient of a scalar function $\phi = \phi(x_1, x_2, x_3)$ is

$$\text{grad}\,\phi = \boldsymbol{\nabla}\phi = \mathbf{e}_i \nabla_i \phi = \mathbf{e}_1 \frac{\partial \phi}{\partial x_1} + \mathbf{e}_2 \frac{\partial \phi}{\partial x_2} + \mathbf{e}_3 \frac{\partial \phi}{\partial x_3}. \tag{1.61}$$

We note that the quantities $\partial x_i / \partial x_s$ in Example 5 above can be written

$$\partial x_i / \partial x_s = \nabla_s x_i = \delta_{is}. \tag{1.62}$$

The divergence of a vector function $\mathbf{a}(x_1, x_2, x_3)$ is

$$\text{div}\,\mathbf{a} = \boldsymbol{\nabla} \cdot \mathbf{a} = \nabla_i a_i = \frac{\partial a_1}{\partial x_1} + \frac{\partial a_2}{\partial x_2} + \frac{\partial a_3}{\partial x_3}. \tag{1.63}$$

The curl of a vector function $\mathbf{a}(x_1, x_2, x_3)$ can be expressed, using (1.38), as

$$(\boldsymbol{\nabla} \times \mathbf{a})_i = \epsilon_{ijk} \nabla_j a_k = \epsilon_{ijk} \frac{\partial a_k}{\partial x_j}, \tag{1.64}$$

giving

$$\text{curl}\,\mathbf{a} = \boldsymbol{\nabla} \times \mathbf{a} = \mathbf{e}_i \epsilon_{ijk} \nabla_j a_k. \tag{1.65}$$

The curl of a vector (in cartesian coordinates) can easily be written down in terms of determinants as follows: if in (1.41) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is written out in full using the definition (1.31) of ϵ_{ijk}, we find

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \epsilon_{ijk} a_i b_j c_k \tag{1.66}$$

$$= a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1, \tag{1.67}$$

which can be written as the determinant

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{1.68}$$

Hence

$$\boldsymbol{\nabla} \times \mathbf{a} = \epsilon_{ijk} \mathbf{e}_i \nabla_j a_k \tag{1.69}$$

$$= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \nabla_1 & \nabla_2 & \nabla_3 \\ a_1 & a_2 & a_3 \end{vmatrix}. \tag{1.70}$$

Various identities involving the vector operator can be derived using the above results (and the summation convention). We illustrate this with three examples.

Example 7

$$\operatorname{div}(\mathbf{b}\phi) = \nabla_i(b_i\phi) \tag{1.71}$$

$$= \phi\nabla_i b_i + (b_i\nabla_i\phi) \tag{1.72}$$

$$= \phi \operatorname{div}\mathbf{b} + (\mathbf{b}\,.\,\boldsymbol{\nabla})\phi. \tag{1.73}$$

Example 8

$$\operatorname{curl}(\mathbf{b}\phi) = \mathbf{e}_i\epsilon_{ijk}\nabla_j(b_k\phi) \tag{1.74}$$

$$= \phi\mathbf{e}_i(\epsilon_{ijk}\nabla_j b_k) + \mathbf{e}_i\epsilon_{ijk}(\nabla_j\phi)b_k \tag{1.75}$$

$$= \phi \operatorname{curl}\mathbf{b} + \mathbf{e}_i\epsilon_{ijk}(\boldsymbol{\nabla}\phi)_j b_k \tag{1.76}$$

$$= \phi \operatorname{curl}\mathbf{b} + (\boldsymbol{\nabla}\phi)\times\mathbf{b}, \tag{1.77}$$

where in the last term we have used (1.38) with $\mathbf{a} = \boldsymbol{\nabla}\phi$.

Example 9

$$\operatorname{curl}\operatorname{curl}\mathbf{A} = \mathbf{e}_i\epsilon_{ijk}\nabla_j(\operatorname{curl}\mathbf{A})_k \tag{1.78}$$

$$= \mathbf{e}_i\epsilon_{ijk}\nabla_j(\epsilon_{klm}\nabla_l A_m) \tag{1.79}$$

$$= \mathbf{e}_i\epsilon_{ijk}\epsilon_{klm}\nabla_l\nabla_j A_m \tag{1.80}$$

$$= \mathbf{e}_i\epsilon_{ijk}\epsilon_{lmk}\nabla_l\nabla_j A_m \tag{1.81}$$

$$= \mathbf{e}_i(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\nabla_l\nabla_j A_m \tag{1.82}$$

$$= \mathbf{e}_i\nabla_i(\nabla_m A_m) - \nabla_j\nabla_j\mathbf{e}_i A_i \tag{1.83}$$

$$= \operatorname{grad}(\operatorname{div}\mathbf{A}) - \nabla^2\mathbf{A}, \tag{1.84}$$

where

$$\nabla^2 = \nabla_j\nabla_j = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

1.6 Orthogonal coordinate systems

So far we have considered only cartesian coordinates x_1, x_2 and x_3. We will require, in later chapters, expressions for the operator div grad $(=\nabla^2)$ in coordinate systems based on cylindrical and spherical polar coordinates.

Consider two points with cartesian coordinates (x_1, x_2, x_3) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)$, where dx_i are small displacements. The infinitesimal distance ds between these two points is given by

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 = dx_i \, dx_i \tag{1.85}$$

(using the summation convention). We now transform to a new coordinate system, say q_1, q_2 and q_3, the x_i being functions of the q_i. If ds^2 can be written in the form

$$ds^2 = h_1^2 \, dq_1^2 + h_2^2 \, dq_2^2 + h_3^2 \, dq_3^2 = h_i^2 \, dq_i^2, \tag{1.86}$$

then the new coordinates form an orthogonal coordinate system. For cartesian coordinates $h_1 = h_2 = h_3 = 1$ and $q_i = x_i$ $(i = 1, 2, 3)$. We now give, without proof, expressions for the gradient, divergence and curl in the new coordinate system in terms of the quantities h_i.

If Φ is a *scalar* and $\mathbf{A} = \mathbf{e}_i A_i$ is a vector then

$$\text{grad } \Phi = \nabla\Phi = \frac{\mathbf{e}_1}{h_1}\frac{\partial \Phi}{\partial q_1} + \frac{\mathbf{e}_2}{h_2}\frac{\partial \Phi}{\partial q_2} + \frac{\mathbf{e}_3}{h_3}\frac{\partial \Phi}{\partial q_3}, \tag{1.87}$$

$$\text{div } \mathbf{A} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}(h_2 h_3 A_1) + \frac{\partial}{\partial q_2}(h_1 h_3 A_2) + \frac{\partial}{\partial q_3}(h_1 h_2 A_3)\right], \tag{1.88}$$

$$\text{curl } \mathbf{A} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1 \mathbf{e}_1 & h_2 \mathbf{e}_2 & h_3 \mathbf{e}_3 \\ \dfrac{\partial}{\partial q_1} & \dfrac{\partial}{\partial q_2} & \dfrac{\partial}{\partial q_3} \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix}. \tag{1.89}$$

It is important to realise that in the above expressions the vectors $\mathbf{e}_i$ are unit vectors which are directed along the three new coordinate axes and point in the direction of increasing coordinate values. Further

$$\nabla^2\Phi = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial \Phi}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_1 h_3}{h_2}\frac{\partial \Phi}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial \Phi}{\partial q_3}\right)\right]. \tag{1.90}$$

We now specialise these results to two particular coordinate systems which are of importance in later chapters.

1. Cylindrical polar coordinates

We specify z, the distance of the point from the x_1, x_2 plane, and the polar coordinates ρ and ϕ of the projection of the point in this plane. Thus $q_1 = \rho$, $q_2 = \phi$, $q_3 = z$ (see Figure 1.2). The vectors $\mathbf{e}_i$ point in the directions of increasing ρ, ϕ and z. The relationships between