Appendix 2

Vectors and tensors

A2.1 COORDINATE SYSTEMS AND TRANSFORMATIONS

The position of a point P in a three-dimensional Euclidean space can be specified by its coordinates (x, y, z) in a rectangular cartesian coordinate system $Oxyz$ with origin O and three mutually perpendicular axes Ox, Oy and Oz. The distance OP from the origin to the point P is given by the three-dimensional generalization of Pythagoras' theorem in plane Euclidean geometry

$$OP = (x^2+y^2+z^2)^{1/2}$$

If n_x, n_y and n_z are the cosines of the angles made by OP with the coordinate axes Ox, Oy and Oz respectively, then

$$x = n_x OP \qquad y = n_y OP \qquad z = n_z OP$$

Thus the rectangular coordinates of any point Q on a line drawn through O and P are proportional to the length OQ.

The choice of coordinate system is initially arbitrary, that is any other rectangular cartesian system would serve as well as the system $Oxyz$. Suppose $Ox'y'z'$ is such another system with the same origin O but mutually perpendicular axes Ox', Oy' and Oz', and suppose the point P has coordinates (x', y', z') in this system. Then by the argument already used, the primed coordinates of any point Q on a line drawn through O and P are also proportional to the length OQ. If Q is such that $OQ = kOP$ then the unprimed and primed coordinates of Q will be (kx, ky, kz) and (kx', ky', kz'), respectively. This implies a linear relationship between the primed and unprimed coordinates that can be written

$$x = a_{xx'}x' + a_{xy'}y' + a_{xz'}z'$$
$$y = a_{yx'}x' + a_{yy'}y' + a_{yz'}z'$$
$$z = a_{zx'}x' + a_{zy'}y' + a_{zz'}z'$$

This and similar relations can be written in much more compact form by relabelling the coordinate axes Ox_1, Ox_2 and Ox_3 instead of Ox, Oy and Oz, and writing the coordinates of P as (x_1, x_2, x_3) instead of (x, y, z). Then the relation between primed and unprimed coordinates can simply be written

$$x_i = a_{ik}x_k' \tag{A2.1}$$

where it is understood that the suffices i, k have the range 1, 2, 3 and that a repeated suffix implies summation over the range of the suffix (Einstein summation convention). Thus, when written out in full,

$$x_i = a_{i1}x_1' + a_{i2}x_2' + a_{i3}x_3' \tag{A2.2}$$

In this notation the square of the length OP is simply written

$$OP^2 = x_ix_i = x_i'x_i' \tag{A2.3}$$

since the length is independent of the coordinate system. This implies a set of conditions on the coefficients a_{ik}, namely

$$a_{ik}a_{il} = \delta_{kl}(= 1 \text{ if } k = l, = 0 \text{ if } k <> l) \tag{A2.4}$$

where δ_{kl} is the Kronecker delta. Using these conditions the equations (A2.1) can be inverted to give the primed coordinates in terms of the unprimed

$$x_i' = a_{ki}x_k \tag{A2.5}$$

In general a coordinate transformation described by Eqns (A2.1) and (A2.4) is termed an orthogonal transformation and is physically represented by either a rigid rotation of the coordinate system, or a rotation together with an inversion, that is a reversal in sense of one of the axes leading to a change from a right-handed to a left-handed system or vice-versa.

A2.2 SCALARS, VECTORS AND TENSORS

A *scalar* quantity is one completely specified by its measure in some previously defined unit. A scalar quantity such as mass, charge or temperature is *invariant* under coordinate transformations.

The physical displacement OP is the prototype *vector* quantity, having both magnitude and direction and being completely specified by the three coordinates x_i of P relative to O in a rectangular coordinate system Ox_i. The x_i are termed the *components* of the displacement vector OP and under an orthogonal transformation of coordinates the components transform according to Eqn (A2.1). In general a vector quantity is defined as any set of three numbers v_i associated with a rectangular cartesian coordinate system Ox_i which transform under an orthogonal transformation in the same way as the components of a physical displacement. The v_i are the *components* of the vector in the coordinate system Ox_i. If v_i' are the components in another coordinate system Ox_i', then *by definition*

$$v_i' = a_{ki}v_k \tag{A2.6}$$

It follows from the kinematic definitions of velocity and acceleration as time derivatives of displacement that they are vector quantities. The product of a vector and a scalar is also a vector quantity, hence momentum and force are also vectors. Since electric charge is a scalar, then it also follows that the electric field is a vector.

If v_i and w_i are two vectors, their *scalar product* or *dot product* is defined as the sum $v_i w_i$. From Eqns (A2.4) and (A2.6) the scalar nature of the dot product follows immediately

$$v_i w_i = a_{ik} a_{il} v_k' w_l' = \delta_{kl} v_k w_l' = v_k' w_k' \tag{A2.7}$$

The *tensor product* of the vectors v_i and w_i is defined as the set of nine quantities $v_i w_k$. The transformation law of the tensor product follows directly from Eqn (A2.6) and is

$$v_i' w_k' = a_{ji} a_{lk} v_j w_l \tag{A2.8}$$

This rule is taken to be the defining property of a *second rank tensor* quantity. Thus a second rank tensor is any collection of nine quantities t_{ik} associated with a rectangular cartesian coordinate system which transform under an orthogonal coordinate transformation according to

$$t_{ik}' = a_{ji} a_{lk} t_{jl} \tag{A2.9}$$

The t_{ik} are the components of the tensor in the coordinate system Ox_i.

Higher rank tensors are defined by straightforward generalization of Eqn (A2.9). Thus a third rank tensor has $3^3 = 27$ components t_{ikm} and a fourth rank tensor $3^4 = 81$ components t_{ikmp}, with transformation laws given by Eqns (A2.10) and (A2.11)

$$t_{ikm}' = a_{ji} a_{lk} a_{nm} t_{jln} \tag{A2.10}$$

$$t_{ikmp}' = a_{ji} a_{lk} a_{nm} a_{qp} t_{jlnq} \tag{A2.11}$$

Any second rank tensor t_{ik} can be written as the sum of a symmetric and an antisymmetric part

$$t_{ik} = t_{(ik)} + t_{[ik]}$$

where

$$t_{(ik)} = (t_{ik} + t_{ki})/2$$
$$t_{[ik]} = (t_{ik} - t_{ki})/2$$

For the symmetric part, $t_{(ik))} = t_{(ki)}$, whereas for the antisymmetric part $t_{[ik]} = -t_{[ki]}$. The antisymmetric part thus has just three independent components $t_{[23]}$, $t_{[31]}$, and $t_{[12]}$. These can be written

$$v_i = \tfrac{1}{2}\epsilon_{ikm} t_{[km]} = \tfrac{1}{2}\epsilon_{ikm} t_{km} \tag{A2.12}$$

where ϵ_{ikm} is the permutation symbol, equal to $+1$ if *ikm* is an even permutation of 123, -1 if ikm is an odd permutation of 123, and 0 if any two of i, k, m are equal. The transformation properties of the v_i can be determined by writing the corresponding expression in the primed coordinate system

$$v_j' = \tfrac{1}{2}\epsilon_{jln}\, t_{[ln]}' = \tfrac{1}{2}\epsilon_{jln}\, a_{kl}a_{mn}t_{[km]}$$

Therefore

$$a_{ij}v_j' = \tfrac{1}{2}\epsilon_{jln}\, a_{ij}a_{kl}a_{mn}t_{[km]}$$

But from the definition of a determinant

$$\epsilon_{ikm}\det\{a_{pq}\} = \epsilon_{jln}\, a_{ij}a_{kl}a_{mn}$$

so that

$$\begin{aligned} v_i &= \det{}^{-1}\{a_{pq}\}\, a_{ij}v_j' \\ v_j' &= \det\{a_{pq}\}\, a_{ij}v_i \end{aligned} \tag{A2.13}$$

Comparing Eqns (A2.13) with (A2.6) shows that the v_i transform as a vector except for the factor $\det\{a_{pq}\}$. From Eqn (A2.4) it follows that the determinant can only have the values $+1$ or -1, since its square must be $+1$. It is also clear that since the determinant is a continuous function of the a_{pq} and must have value $+1$ for the identity transformation $a_{pq} = \delta_{pq}$, it must have the value $+1$ for all transformations obtained by a continuous variation of the a_{pq} from δ_{pq}. Physically, this means that all orthogonal transformations corresponding to a rotation of the coordinate system have $\det\{a_{pq}\} = +1$, whereas transformations that involve a change in sense of the coordinate system have $\det\{a_{pq}\} = -1$. The former transformations are termed *proper*, the latter *improper*.

Equations (A2.13) then show that the v_i transform as a vector under proper orthogonal transformations, but change sign relative to a true vector under improper transformations. Generally, Eqns (A2.13) are used as the defining property for a class of physical quantities known as *axial* or *pseudo-* vectors. The mathematical term for quantities such as v_i is a *relative vector of weight +1*. By extension, a relative tensor of weight W obeys by definition a transformation law

$$t_{jln..}' = \det\{a_{pq}\}^W a_{ij}a_{kl}a_{mn}..\, t_{ikm..} \tag{A2.14}$$

Of particular interest is the relative vector u_i associated with the antisymmetric part of the tensor product of two vectors v_i and w_i

$$u_i = \epsilon_{ikm}\, v_{[k}w_{m]} = \epsilon_{ikm}\, v_k w_m \tag{A2.15}$$

This is the *vector product* or *cross product* of v_i and w_i. Equation (A2.15) shows clearly that under an inversion of the coordinate axes in which the components of a true or absolute vector change sign, the components of the axial vector u_i are unaltered.

A2.3 LENGTHS, ANGLES, AREAS AND VOLUMES

Suppose that $x_i^{(1)}$ are the components of the displacement vector from the origin O to a point $P^{(1)}$. Similarly, let $x_i^{(2)}$, $x_i^{(3)}$ be the components of the displacements from O to points $P^{(2)}$ and $P^{(3)}$, respectively. Then the length or magnitude L of the displacement $OP^{(1)}$ is given by

$$L^2 = \delta_{kl} x_k^{(1)} x_l^{(1)} = x_k^{(1)} x_k^{(1)} \tag{A2.16}$$

If the angle between $OP^{(1)}$ and $OP^{(2)}$ is β and the lengths of $OP^{(1)}$ and $OP^{(2)}$ are $L_{(1)}$ and $L_{(2)}$, respectively, then

$$x_i^{(1)} x_i^{(2)} = L_{(1)} L_{(2)} \cos(\beta)$$

In particular, if $OP^{(1)}$ and $OP^{(2)}$ are perpendicular, then $x_i^{(1)} x_i^{(2)} = 0$.

Now consider the parallelogram with adjacent edges $OP^{(1)}$ and $OP^{(2)}$. Let A_{23} be the area of the projection of the parallelogram on the Ox_2x_3 coordinate plane. Then A_{23} is given by

$$A_{23} = x_2^{(1)} x_3^{(2)} - x_3^{(1)} x_2^{(2)} = 2x_{[2}^{(1)} x_{3]}^{(2)}$$

Cyclic permutation of the indices gives the projections A_{31} and A_{12} on the remaining coordinate planes. The quantities A_{ik} form the components of an antisymmetric tensor

$$A_{ik} = 2x_{[i}^{(1)} x_{k]}^{(2)}$$

In analogy with Eqn (A2.12), a relative vector of weight $+1$ can be associated with A_{ik} by

$$A_i = \tfrac{1}{2}\epsilon_{ikm} A_{km} = \epsilon_{ikm} x_k^{(1)} x_m^{(2)} \tag{A2.17}$$

A_i is the vector area of the parallelogram with adjacent edges $OP^{(1)}$ and $OP^{(2)}$. From Eqn (A2.17) it follows that $A_i x_i^{(1)} = A_i x_i^{(2)} = 0$, so that A_i is perpendicular to both $x_i^{(1)}$ and $x_i^{(2)}$. It also follows that the sign of A_i depends on the order in which the edge vectors are taken, that is, the area has two definite orientations.

Finally, consider the parallelepiped with adjacent edges $OP^{(1)}$, $OP^{(2)}$ and $OP^{(3)}$. The volume V of the parallelepiped is given by the scalar product of one of the edge vectors with the vector area of the parallelogram defined by the other two. Therefore

$$V = x_i^{(1)} \epsilon_{ikm} x_k^{(2)} x_m^{(3)} = \epsilon_{ikm} x_i^{(1)} x_k^{(2)} x_m^{(3)} \tag{A2.18}$$

Again the sign of V depends on the order in which the edge vectors are taken, so that V has two possible orientations. Under an orthogonal transformation, V transforms according to

$$V' = \epsilon_{jln} x_j^{(1)\prime} x_l^{(2)\prime} x_n^{(3)\prime} = \epsilon_{jln} a_{ij} a_{kl} a_{mn} x_i^{(1)} x_k^{(2)} x_m^{(3)}$$

Therefore

$$V' = \det\{a_{pq}\}\,\epsilon_{ikm}x_i^{(1)}x_k^{(2)}x_m^{(3)} = \det\{a_{pq}\}\,V \tag{A2.19}$$

V is thus a relative scalar of weight $+1$.

A2.4 VECTOR AND TENSOR FIELDS

The mass or charge associated with a particle depends only on the particle and, at least in classical mechanics, is independent of the position of the particle. Other physical quantities are however defined at each point in a region of space and over an interval of time and have to be expressed mathematically as functions of position and time. Such quantities are known as *fields*. A *scalar field* F is represented by a single function of position x_i and time t, so that $F = F(x_i, t)$. By definition the value of the field F at a given point is invariant under coordinate transformations, although the form of its functional dependence on the coordinates will be course depend on the coordinate system chosen. Thus if $x_i = a_{ik}x_k'$ is an orthogonal transformation

$$F = F(x_i, t) = F(a_{ik}x_k', t) = F'(x_i', t)$$

An example of a scalar field is the temperature at each point of a continuum, $T = T(x_i, t)$.

Vector and tensor fields are by extension sets of functions $t_{ik..}(x_r, t)$ defined over a given space–time region in such a way that the values of the functions at a given space–time point transform according to the appropriate law under a coordinate transformation

$$t_{ik..}(x_r, t) = a_{ij}a_{kl}\cdot\,.\,t_{jl..}'(x_r, t)$$

Clearly both the value and the functional form of a component of a vector or tensor field depends on the coordinate system chosen.

Let F be a scalar field $F = F(x_i, t)$. Define $F_{,i}$ to be the partial derivative of F with respect to x_i. Then the $F_{,i}$ form the components of a vector field known as the ***gradient*** of the scalar field F. This follows from the elementary rules of calculus

$$F_{,i} = F_{,k}'x_k'_{,i} = a_{ik}F_{,k}' \tag{A2.20}$$

where $F_{,k}'$ is the partial derivative of F with respect to x_k'. Equation (A2.20) shows that the gradient $F_{,i}$ satisfies the defining rule (A2.6) for a vector quantity.

Now let F_i be a vector field. The sum $F_{i,i} = F_{1,1} + F_{2,2} + F_{3,3}$ is a scalar field known as the ***divergence*** of F_i. Its scalar nature follows directly from

$$F_{i,i} = (a_{ik}F_k')_{,j}'x_j'_{,j} = a_{ik}a_{ij}F_k'_{,j} = F_k'_{,k}'$$

where the orthogonality condition Eqn (A2.4) has been used.

In general the partial derivatives $F_{i,k}$ of a vector field form the components of a second rank tensor. The antisymmetric part $F_{[i,k]}$ has just three independent components $F_{[2,3]}$, $F_{[3,1]}$ and $F_{[1,2]}$, which can be written

$$v_i = \epsilon_{ikm} F_{[k,m]} = \epsilon_{ikm} F_{k,m} \tag{A2.21}$$

The v_i transform as a relative vector of weight 1 known as the *curl* of the vector field F_i.

The importance of the divergence and curl of a vector field F_i lies in their appearance in the integral theorems of Gauss and Stokes. First let V be a region in space bounded by the closed surface S. At each point of S let n_i be the outward-pointing unit normal (ie, a vector of unit length drawn perpendicular to S and pointing out of V). Then if F_i is a vector field defined over V, Gauss's theorem states that

$$\int_V F_{i,i}\,\mathrm{d}V = \oint_S F_i n_i\,\mathrm{d}S \tag{A2.22}$$

The restriction of F_i to a vector field is not necessary, and Gauss's theorem may be stated for any tensor field $t_{ik..p}$

$$\int_V t_{ik..p,p}\,\mathrm{d}V = \oint_S t_{ik..p} n_p\,\mathrm{d}S \tag{A2.23}$$

Gauss's theorem is frequently referred to as the divergence theorem.

If now S is a surface bounded by a closed loop C, and if the unit normal n_i on S and the positive sense of C form a right-handed screw, Stokes's theorem states that

$$\int_S \epsilon_{ikm}\, n_i F_{k,m}\,\mathrm{d}S = \oint_C F_i\,\mathrm{d}l_i \tag{A2.24}$$

where the $\mathrm{d}l_i$ are the components of an element of length along the boundary curve C.

A2.5 CHANGE OF VARIABLES IN A MULTIPLE INTEGRAL

In continuum mechanics it is frequently convenient in applications of the divergence theorem to make use of the rule for change of variables in a multiple integral. For completeness, this rule is stated here.

Let f be a function of the n real variables x_i, where i now runs from 1 to n. The x_i can be regarded as the coordinates of a point in an n-dimensional space, x-space. Let X_i be a second set of n variables, which can similarly be regarded as the coordinates of a point in an n-dimensional X-space. Suppose now that the x_i and the X_i are connected by a set of functions $x_i = x_i(X_k)$ and their inverses $X_i = X_i(x_k)$. The necessary and sufficient condition that the inverses should exist is that the *Jacobian determinant* J should not vanish in the region of interest

$$\mathrm{J} = \det\{x_{i,k}\} = 1/\det\{X_{i,k}\} \neq 0$$

If f is defined in a region V of x-space and V' is the region in X-space into which V is mapped by the functions $X_i(x_k)$, then the integral of f over V can be transformed into an integral over V' by

$$\int_V \qquad \int_{V'} \qquad)\ \mathrm{J}\ \mathrm{d}X_1 \mathrm{d}X_2 .\ .\mathrm{d}X_n$$

or

$$\int_V f(x_i)\mathrm{d}V = \int_{V'} f(x_i(X_k))\mathrm{J}\mathrm{d}V' \qquad \text{(A2.25)}$$

A2.6 EIGENVALUES OF A REAL SYMMETRIC TENSOR

For any second rank tensor t with components t_{km} and any vector v with components v_m, the products $w_k = t_{km}v_m$ form the components of a second vector w, which can be described as the result of the tensor t *operating* on the vector v. The *eigenvalue problem* for the tensor t is the problem of determining all vectors e and scalars E such that

$$t_{km}e_m = Ee_k \qquad \text{(A2.26)}$$

th : effect of t operating on e is to leave the direction of e
ur e satisfying Eqn (A2.26) are *eigenvectors* of t, and the
sc responding *eigenvalues*.
can be written as

$$(t_{km} - E\delta_{km})e_m = 0 \qquad \text{(A2.27)}$$

which is a system of homogeneous linear equations for the components e_m. The condition for non-trivial solutions to exist is that the determinant of coefficients should vanish

$$\det\{t_{km} - E\delta_{km}\} = 0$$

Expansion of the determinant leads to a cubic equation in E, which in general will have three (complex) roots $E^{(1)}$, $E^{(2)}$ and $E^{(3)}$. For each eigenvalue $E^{(k)}$, Eqns (A2.27) can be solved for the (complex) components $e_m^{(k)}$ of the associated eigenvector. (Note that since the eigenvectors are only determined up to a scalar multiple, they can be assumed to be normalized to unit length through $e_m e_m^* = 1$, where the * indicates the complex conjugate.)

Provided the components t_{km} of t are real, the coefficients of the cubic in E will all be real, and it then follows that there will be at least one real eigenvalue, and an associated real eigenvector. If in addition to being real, the t_{km} are also symmetric, so that $t_{km} = t_{mk}$, then it can be proved that all the eigenvalues and eigenvectors are real. For if E and e_m satisfy

$$t_{km}e_m = Ee_k \tag{A2.28}$$

then by taking complex conjugates it follows that

$$t_{km}e_m^* = E^*e_k^* \tag{A2.29}$$

Multiplying Eqns (A2.28) and (A2.29) by e^*_k and e_k, respectively, and subtracting leads to

$$e_k^*t_{km}e_m - e_kt_{km}e_m^* = (E-E^*)e_ke_k^* \tag{A2.30}$$

But by the symmetry of t_{km} the left side of Eqn (A2.30) vanishes and hence $E=E^*$, that is, E is real. It then follows from Eqn (A2.27) that the associated eigenvector e_k is also real.

If now E and e_k, F and f_k are two eigenvalues and associated eigenvectors, it follows that

$$f_kt_{km}e_m = Ef_ke_k$$
$$e_kt_{km}f_m = Ff_ke_k$$

Again by the symmetry of t_{km} the left sides are equal so that subtracting leads to

$$(E-F)f_ke_k = 0 \tag{A2.31}$$

Hence if E and F are not equal it follows that the associated eigenvectors are orthogonal. If $E=F$ this does not follow, but in such a case it is easily shown that any linear combination of e_k and f_k is also an eigenvector of t_{km} corresponding to the same eigenvalue $E=F$. Thus it is always possible to choose eigenvectors such that $e_kf_k=0$.

In summary a real symmetric tensor t has three real eigenvalues. In the case where the eigenvalues are distinct, the associated eigenvectors are necessarily mutually orthogonal, whereas in the degenerate case where two or all three eigenvalues are equal, the eigenvectors may be chosen to be orthogonal.

Supposing that three orthonormal eigenvectors $e_m^{(k)}$ have been selected, it is then possible to use these as the basis for a new coordinate system, that is a coordinate system Ox_k' in which the eigenvectors have components (1, 0, 0), (0, 1, 0) and (0, 0, 1). If the coordinate transformation is described by $x_k = a_{km}x_m'$ as in Eqn (A2.1), then clearly $e_m^{(k)} = a_{mk}$. The components of t in the new coordinate system are given via Eqns (A2.9) as

$$t_{ik}' = a_{ji}a_{lk}t_{jl} = e_j^{(i)}e_l^{(k)}t_{jl}$$

But since the $e_m^{(k)}$ are eigenvectors, this reduces to

$$t_{ik}' = E^{(k)}\delta_{ik} \qquad \text{(no sum on } k) \tag{A2.32}$$

That is, in the coordinate system defined by the eigenvectors, the components t_{km}' vanish if $k <> m$, and the diagonal entries t_{11}', t_{22}' and t_{33}' are just the eigenvalues $E^{(1)}$, $E^{(2)}$ and $E^{(3)}$.

Finally, consider the scalar quantity $Q = t_{km}x_kx_m$. In the eigenvector coordinate system, Q takes the simple form

$$Q = E^{(1)}x_1^2 + E^{(2)}x_2^2 + E^{(3)}x_3^2$$

If all the eigenvalues are positive, it then follows that $Q > 0$ for all values of x_k except $x_k = 0$, when $Q = 0$. When this is the case, the real symmetric tensor t is described as *positive-definite*. Reversing the argument, if it can be shown that $Q > 0$ for all x_k not equal to zero, then it follows that all the eigenvalues of the associated real, symmetric tensor are real and *positive*. (This result is of fundamental importance in the theory of wave propagation in anisotropic materials, discussed in Chapter 2.)