CHAPTER 1

On Vectors and Tensors, Expressed in Cartesian Coordinates

It's not enough, to characterize a vector as "something that has magnitude and direction."

First, we'll look at something with magnitude and direction, that is *not* a vector.

Second, we'll look at a similar example of something that *is* a vector, and we'll explore some of its properties.

Then we'll give a formal definition of a cartesian vector — that is, a vector whose components we choose to analyse in cartesian coordinates. The definition can easily be generalized to cartesian tensors.

Finally on this subject, we'll explore some basic properties of cartesian tensors, showing how they extend the properties of vectors. As examples, we'll use the strain tensor, the stress tensor, and (briefly) the inertia tensor.

Figure 1.1 shows the outcome of a couple of rotations, applied first in one sequence, then in the other sequence.

We see that if we add the second rotation to the first rotation, the result is different from adding the first rotation to the second rotation.

So, finite rotations do not commute. They each have magnitude (the angle through which the object is rotated), and direction (the axis of rotation). But

First rotation + Second rotation $\neq$ Second rotation + First rotation.

However, *infinitesimal rotations*, and *angular velocity*, truly are vectors. To make this point, we can use intuitive ideas about displacement (which is a vector). Later, we'll come back to the formal definition of a vector.

What is angular velocity? We define $\boldsymbol{\omega}$ as rotation about an axis (defined by unit vector $\mathbf{l}$, say) with angular rate $\frac{d\Omega}{dt}$, where Ω is the angle through which the line PQ (see Figure 1.2) moves with respect to some reference position (the position of PQ at the reference time). So

$$\boldsymbol{\omega} = \frac{d\Omega}{dt}\mathbf{l},$$

and Ω = finite angle = $\Omega(t)$.

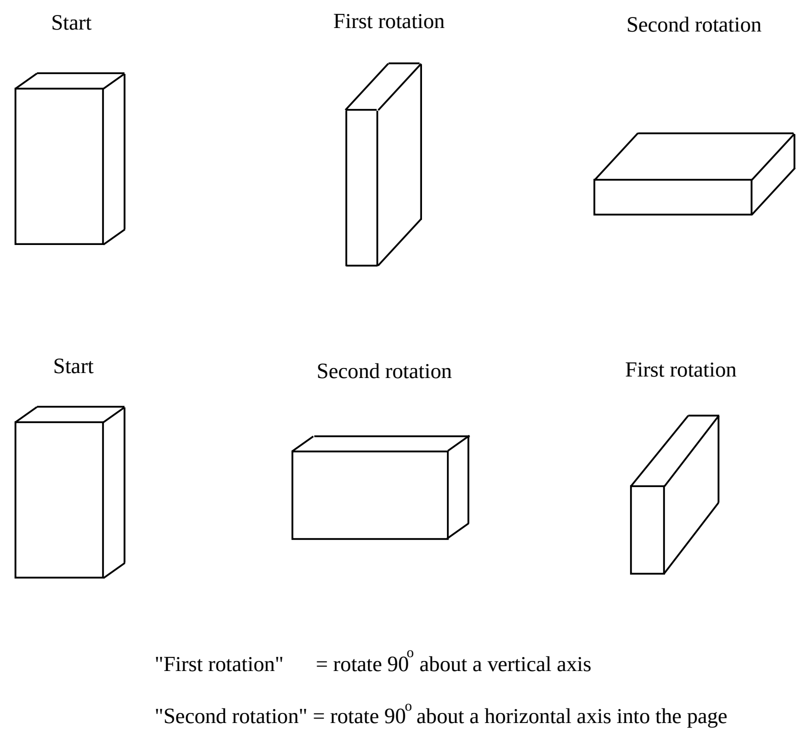

FIGURE **1.1**
A block is shown after various rotations. Applying the first rotation and then the second, gives a different result from applying the second rotation and then the first.

To prove that angular velocity $\boldsymbol{\omega}$ is a vector, we begin by noting that the infinitesimal rotation in time dt is $\boldsymbol{\omega} dt = \mathbf{l} d\Omega$. During the time interval dt (think of this as δt, then allow $\delta t \to 0$), the displacement of P has amplitude QP times $d\Omega$. This amplitude is $r \sin\theta d\Omega$, and the direction is perpendicular to $\mathbf{r}$ and $\mathbf{l}$, i.e.

$$\text{displacement} = (\boldsymbol{\omega} \times \mathbf{r}) dt.$$

Displacements add vectorially. Consider two simultaneous angular velocities $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$. Then the total displacement (of particle P in time dt) is

$$(\boldsymbol{\omega}_1 \times \mathbf{r}) dt + (\boldsymbol{\omega}_2 \times \mathbf{r}) dt = (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \times \mathbf{r} dt$$

for all $\mathbf{r}$. We can interpret the right-hand side as a statement that the total angular velocity is $\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2$. If we reversed the order, then the angular velocity would be the same sum, $\boldsymbol{\omega}_2 + \boldsymbol{\omega}_1$, associated with the same displacement, so $\boldsymbol{\omega}_2 + \boldsymbol{\omega}_1 = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2$.

So: sometimes entities with magnitude and direction obey the basic commutative rule that $\mathbf{A}_2 + \mathbf{A}_1 = \mathbf{A}_1 + \mathbf{A}_2$, and sometimes they do not.

What then is a vector? It is an entity that in practice is studied quantitatively in terms

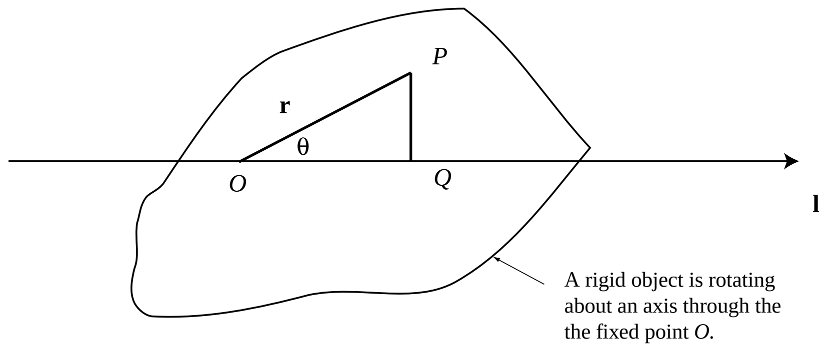

FIGURE **1.2**
P is a point fixed in a rigid body that rotates with angular velocity $\boldsymbol{\omega}$ about an axis through O. The point Q lies at the foot of the perpendicular from P onto the rotation axis.

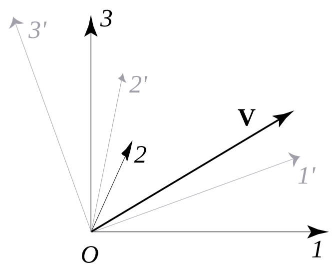

FIGURE **1.3**
Here is shown a vector **V** together with an original cartesian coordinate system having axes $Ox_1x_2x_3$ (abbreviated to $O1$, $O2$, $O3$). Also shown is another cartesian coordinate system with the same origin, having axes $Ox_1'x_2'x_3'$. Each system is a set of mutually orthogonal axes.

of its components. In cartesians a vector **V** is expressed in terms of its components by

$$\mathbf{V} = V_1\hat{\mathbf{x}}_1 + V_2\hat{\mathbf{x}}_2 + V_3\hat{\mathbf{x}}_3 \tag{1.1}$$

where $\hat{\mathbf{x}}_i$ is the unit vector in the direction of the i-axis. An alternative way of writing equation (1.1) is $\mathbf{V} = (V_1, V_2, V_3)$, and sometimes just the symbol V_i. Then $V_1 = \mathbf{V} \cdot \hat{\mathbf{x}}_1$ and in general $V_i = \mathbf{V} \cdot \hat{\mathbf{x}}_i$. Thus, when writing just V_i, we often leave understood (a) the fact that we are considering all three components (i = 1, 2, or 3); and (b) the fact that these particular components are associated with a particular set of cartesian coordinates.

What then is the significance of working with a *different* set of cartesian coordinate axes? We shall have a different set of components of a given vector, **V**. See Figure 1.3 for an illustration of two different cartesian coordinate systems. What then are the new components of **V**?

We now have

$$\mathbf{V} = V_1'\hat{\mathbf{x}}_1' + V_2'\hat{\mathbf{x}}_2' + V_3'\hat{\mathbf{x}}_3'$$

where $\hat{\mathbf{x}}_1'$ is a unit vector in the new x_j' – direction. So the new components are V_j'. Another way to write the last equation is $\mathbf{V} = (V_1', V_2', V_3')$, which is another expression of the same vector $\mathbf{V}$, this time in terms of its components in the new coordinate system.

Then (a third way to state the same idea),

$$V_j' = \mathbf{V} \cdot \hat{\mathbf{x}}_j'. \tag{1.2}$$

We can relate the new components to the old components, by substituting from (1.1) into (1.2), so that

$$V_j' = (V_1\hat{\mathbf{x}}_1 + V_2\hat{\mathbf{x}}_2 + V_1\hat{\mathbf{x}}_3) \cdot \hat{\mathbf{x}}_j' = \sum_{i=1}^{3} l_{ij} V_i$$

where

$$l_{ij} = \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j'.$$

Since l_{ij} is the dot product ot two unit vectors, it is equal to the cosine of the angle between $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j'$; that is, the cosine of the angle between the original x_i–axis and the new x_j'–axis.

The l_{ij} are often called *direction cosines*. In general, $l_{ji} \neq l_{ij}$: they are not symmetric, because l_{ji} is the cosine of the angle between the x_j-axis and the x_i'-axis, and in general this angle is independent of the angle between x_i- and x_j'-axes. But we don't have to be concerned about the order of the axes, in the sense that $\cos(-\theta) = \cos\theta$ so that l_{ij} is also the cosine of the angle between the x_j'–axis and the x_i–axis.

At last we are in a position to make an important definition. We say that $\mathbf{V}$ is a *cartesian vector* if its components V_j' in a new cartesian system are obtained from its components V_i in the previously specified system by the rule

$$V_j' = \sum_{i=1}^{3} l_{ij} V_i. \tag{1.3}$$

This definition indicates that the vector $\mathbf{V}$ has meaning, independent of any cartesian coordinate system. When we express $\mathbf{V}$ in terms of its components, then they will be different in different coordinate systems; and those components transform according to the rule (1.3). This rule is the defining property of a cartesian vector.

It is time now to introduce the Einstein summation convention — which is simple to state, but whose utility can be appreciated only with practice. According to this convention, we don't bother to write the summation for equations such as (1.3) which have a pair of repeated indices. Thus, with this convention, (1.3) is written

$$V_j' = l_{ij} V_i \tag{1.4}$$

and the presence of the "$\sum_{i=1}^{3}$" is flagged by the once-repeated subscript i. Even though we don't bother to write it, we must not forget that this unstated summation is still required over such repeated subscripts.

[Looking ahead, we shall find that the rule (1.4) can be generalized for entities called *second-order cartesian tensors*, symbolized by **A**, with cartesian coordinates that differ in the new and original systems. The defining property of such a tensor is that its components in different coordinate systems obey the relationship

$$A'_{jl} = A_{ik} l_{ij} l_{kl}.]$$

As an example of the summation convention, we can write the scalar product of two vectors **a** and **b** as

$$\mathbf{a} \cdot \mathbf{b} = a_i b_i.$$

The required summation over i in the above equation, is (according to the summation convention) signalled by the repeated subscript. Note that the repeated subscript could be any symbol. For example we could replace i by p, and write $a_i b_i = a_p b_p$. Because it doesn't matter what symbol we use for the repeated subscript in the summation convention, i or p here is called a *dummy subscript*. Any symbol could be used (as long as it is repeated).

The Einstein summation convention is widely used together with symbols δ_{ij} and ε_{ijk} defined as follows:

$$\delta_{ij} = 0 \quad \text{for} \quad i \neq j, \quad \text{and} \quad \delta_{ij} = 1 \quad \text{for } i = j; \tag{1.5}$$

and

$$\begin{aligned} &\varepsilon_{ijk} = 0 \quad \text{if any of } i, j, k \text{ are equal, otherwise} \\ &\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = -\varepsilon_{213} = -\varepsilon_{321} = -\varepsilon_{132} = 1. \end{aligned} \tag{1.6}$$

Note that ε_{ijk} is unchanged in value if we make an even permutation of subscripts (such as $123 \to 312$), and changes sign for an odd permutation (such as $123 \to 213$).

The most important properties of the symbols in (1.5) and (1.6) are then

$$\delta_{ij} a_j = a_i, \qquad \varepsilon_{ijk} a_j b_k = (\mathbf{a} \times \mathbf{b})_i, \tag{1.7}$$

for any vectors **a** and **b**; and the symbols are linked by the property

$$\varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{jl} & \delta_{kl} \\ \delta_{im} & \delta_{jm} & \delta_{km} \\ \delta_{in} & \delta_{jn} & \delta_{kn} \end{vmatrix} \tag{1.8}$$

from which it follows that

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{1.9}$$

To prove (1.8), note that if any pair of (i, j, k) or any pair of (l, m, n) are equal, then the left-hand side and right-hand side are both zero. (A determinant with a pair of equal rows or a pair of equal columns is zero.) If $(i, j, k) = (l, m, n) = (1, 2, 3)$, then the left-hand side and right-hand side are both 1 because $\varepsilon_{123}\varepsilon_{123} = 1$ and

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1. \tag{1.10}$$

Any other example of (1.8) where the left-hand side is not zero, will require that the subscripts (i, j, k) be an even or an odd permutation of $(1, 2, 3)$, and similarly for the subscripts (k, l, m), giving a value (for the left-hand side) equal either to 1 or to -1. But the same type of permutation of (i, j, k) or (l, m, n) (whether even or odd) will also apply to columns or to rows of (1.10), giving either 1 (for a net even permutation) or -1 (for a net odd permutation), and again the left-hand side of (1.8) equals the right-hand side.

Because of the first of the relations given in (1.7), δ_{ij} is sometimes called the *substitution symbol* or *substitution tensor*. In recognition of its originator it also called the *Kronecker delta*. ε_{ijk} is usually called the *alternating tensor*.

(1.9) follows from (1.8), recognizing that we need to allow for the summation over i. Thus

$$\varepsilon_{ijk}\varepsilon_{ilm} = \begin{vmatrix} \delta_{ii} & \delta_{ji} & \delta_{ki} \\ \delta_{il} & \delta_{jl} & \delta_{kl} \\ \delta_{im} & \delta_{jm} & \delta_{km} \end{vmatrix}$$
$$= \delta_{ii}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) - \delta_{ji}(\delta_{il}\delta_{km} - \delta_{im}\delta_{kl}) + \delta_{ki}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}),$$

but here δ_{ii} is not equal to 1 (which is what most people who are unfamiliar with the summation convention might think at first). Rather, $\delta_{ii} = \sum_{i=1}^{3} \delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3$. Using this result, and the "substitution" property of the Kronecker delta function (the first of the relations in (1.7)), we find

$$\varepsilon_{ijk}\varepsilon_{ilm} = 3(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) - (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) + (\delta_{kl}\delta_{jm} - \delta_{km}\delta_{jl}),$$

which simplifies to (1.9) after combining equal terms. As we should expect, the subscript i does not appear in the right-hand side.

1.1 Tensors

Tensors generalize many of the concepts described above for vectors. In this Section we shall look at tensors of stress and strain, showing in each case how they relate a pair of vectors. We shall develop

(i) the physical ideas behind a particular tensor (for example, stress or strain);

(ii) the notation (for example, for the cartesian components of a tensor);

(iii) a way to think conceptually of a tensor, that avoids dependence on any particular choice of coordinate system; and

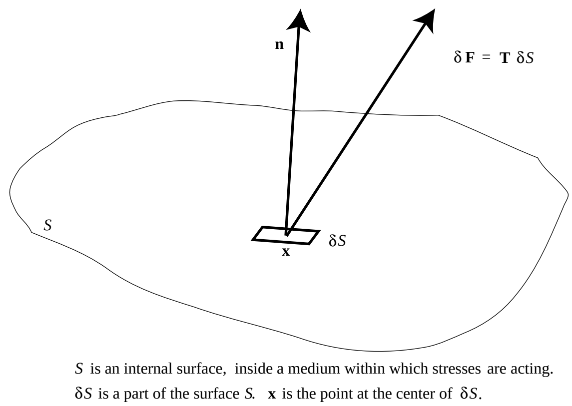

FIGURE **1.4**
The definition of traction $\mathbf{T}$ acting at a point across the internal surface S with normal $\mathbf{n}$ (a unit vector). The choice of sign is such that traction is a pulling force. Pushing is in the opposite direction, so for a fluid medium, the pressure would be $-\mathbf{n} \cdot \mathbf{T}$.

(iv) the formal definition of a tensor (analogous to the definition of a vector based on (1.3) or (1.4)).

To analyze the internal forces acting mutually between adjacent particles within a continuum, we use the concepts of *traction* and *stress tensor*. Traction is a vector, being the force acting per unit area across an internal surface within the continuum, and quantifies the contact force (per unit area) with which particles on one side of the surface act upon particles on the other side. For a given point of the internal surface, traction is defined (see Fig. 1.4) by considering the infinitesimal force $\delta\mathbf{F}$ acting across an infinitesimal area δS of the surface, and taking the limit of $\delta\mathbf{F}/\delta S$ as $\delta S \to 0$. With a unit normal $\mathbf{n}$ to the surface S, the convention is adopted that $\delta\mathbf{F}$ has the direction of force due to material on the side to which $\mathbf{n}$ points and acting upon material on the side from which $\mathbf{n}$ is pointing; the resulting traction is denoted as $\mathbf{T}(\mathbf{n})$. If $\delta\mathbf{F}$ acts in the direction shown in Fig. 1.4, traction is a pulling force, opposite to a pushing force such as pressure. Thus, in a fluid, the (scalar) pressure is $-\mathbf{n} \cdot \mathbf{T}(\mathbf{n})$. For a solid, shearing forces can act across internal surfaces, and so $\mathbf{T}$ need not be parallel to $\mathbf{n}$. Furthermore, the magnitude and direction of traction depend on the orientation of the surface element δS across which contact forces are taken (whereas pressure at a point in a fluid is the same in all directions). To appreciate this orientation-dependence of traction at a point, consider a point P, as shown in Figure 1.5, on the exterior surface of a house. For an element of area on the surface of the wall at P, the traction $\mathbf{T}(\mathbf{n}_1)$ is zero (neglecting atmospheric pressure and winds); but for a horizontal element of area within the wall at P, the traction $\mathbf{T}(\mathbf{n}_2)$ may be large (and negative).

Because $\mathbf{T}$ can vary from place to place, as well as with orientation of the underlying element of area needed to define traction, $\mathbf{T}$ is separately a function of $\mathbf{x}$ and $\mathbf{n}$. So we write $\mathbf{T} = \mathbf{T}(\mathbf{x}, \mathbf{n})$.

At a given position $\mathbf{x}$, the stress tensor is a device that tells us how $\mathbf{T}$ depends upon $\mathbf{n}$. But before we investigate this dependence, we first see what happens if $\mathbf{n}$ changes sign.

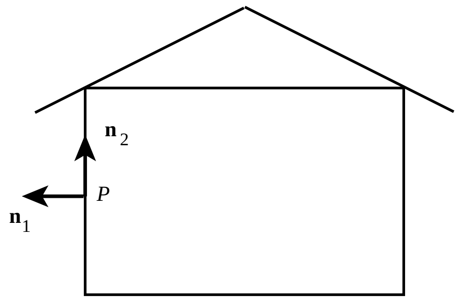

FIGURE **1.5**
$\mathbf{T}(\mathbf{n}_1) \neq \mathbf{T}(\mathbf{n}_2)$. The traction vector in general is different for different orientations of the area across with the traction is acting.

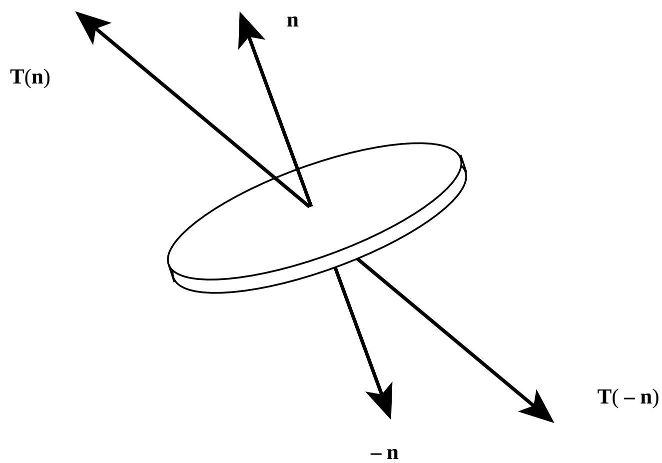

FIGURE **1.6**
A disk with parallel faces. The normals to opposite faces have the same direction but opposite sign.

By considering a small disk-shaped volume (Figure 1.6) whose opposite faces have opposite normals $\mathbf{n}$ and $-\mathbf{n}$, we must have a balance of forces

$$\mathbf{T}(-\mathbf{n}) = -\mathbf{T}(\mathbf{n}) \tag{1.11}$$

otherwise the disk would have infinite acceleration, in the limit as its volume shrinks down to zero. (There is negligible effect from the edges as they are so much smaller than the flat faces.)

In a similar fashion we can examine the balance of forces on a small tetrahedron that has three of its four faces within the planes of a cartesian coordinate system, as shown in Figure 1.7. The oblique (fourth) face of the tetrahedron has normal $\mathbf{n}$ (a unit vector), and by projecting area ABC onto each of the coordinate planes, we find the following relation between areas:

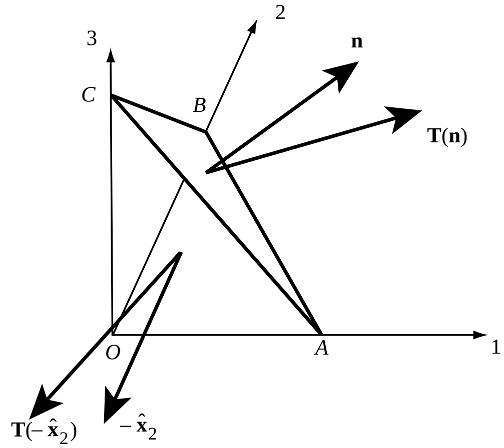

FIGURE **1.7**
The small tetrahedron $OABC$ has three of its faces in the coordinate planes, with outward normals $-\hat{\mathbf{x}}_j (j = 1, 2, 3)$ (only one of which is shown here, $j = 2$), and the fourth face has normal $\mathbf{n}$.

$$(OBC, OCA, OAB) = ABC\ (n_1, n_2, n_3). \tag{1.12}$$

There are four forces acting on the tetrahedron, one for each face. Thus, face OBC has the outward normal given by the unit vector $-\hat{\mathbf{x}}_1 = (-1, 0, 0)$. This face is pulled by the traction $\mathbf{T}(-\hat{\mathbf{x}}_1)$, and hence by the force $\mathbf{T}(-\hat{\mathbf{x}}_1)$ times area OBC (remember, traction is force per unit area). The balance of forces then requires that

$$\mathbf{T}(\mathbf{n})\ ABC + \mathbf{T}(-\hat{\mathbf{x}}_1)\ OBC + \mathbf{T}(-\hat{\mathbf{x}}_2)\ OCA + \mathbf{T}(-\hat{\mathbf{x}}_3)\ OAB = \mathbf{0}. \tag{1.13}$$

(If the right-hand side were not zero, we would get infinite acceleration in the limit as the tetrahedron shrinks down to the point O.) Using the two equations (1.11)–(1.13), it follows that

$$\begin{aligned} \mathbf{T}(\mathbf{n}) &= \mathbf{T}(\hat{\mathbf{x}}_1)n_1 + \mathbf{T}(\hat{\mathbf{x}}_2)n_2 + \mathbf{T}(\hat{\mathbf{x}}_3)n_3 \\ &= \mathbf{T}(\hat{\mathbf{x}}_j)n_j \qquad \text{(using the summation convention).} \end{aligned} \tag{1.14}$$

If we now define

$$\tau_{kl} = T_l(\hat{\mathbf{x}}_k), \tag{1.15}$$

then

$$T_i(\mathbf{n}) = \tau_{ji} n_j. \tag{1.16}$$

If we can show that $\tau_{ji} = \tau_{ij}$, then

$$T_i = \tau_{ij} n_i. \tag{1.17}$$

This equation gives a simple rule by which the components of the traction vector, T_i, are given as a linear combination of the components of the normal vector n_j. The nine symbols τ_{ij} are the cartesian components of a tensor, namely, the *stress tensor*. First, we'll show that indeed $\tau_{ji} = \tau_{ij}$. Second, we'll show that the symbols τ_{ij} specify a surface which is independent of our particular choice of coordinate axes.

1.1.1 SYMMETRY OF THE STRESS TENSOR

To see why the τ_{ij} are symmetric, we can look in some detail at a particular example, namely τ_{21} and τ_{12}. They quantify components of the tractions $\mathbf{T}$ (force per unit area) acting on the faces of a small cube with sides of length δx_1, δx_2, δx_3 as shown in Figure 1.8.

The force acting on the top face of the cube is traction $\times$ area, which is $\mathbf{T}(\hat{\mathbf{x}}_2)\ \delta x_1\ \delta x_2$. And on the opposite face the force is $\mathbf{T}(\widehat{-\mathbf{x}}_2)\ \delta x_1\ \delta x_2$. From these two faces, what is the strength of the resulting couple that tends to make the cube rotate about the x_3 axis? The x_2 and x_3 components of $\mathbf{T}$ have no relevance here (they are associated with the tendency to rotate about different axes) — only the x_1 component of $\mathbf{T}$, which is τ_{21}. Figure 1.8b shown the resulting couple, and it is $\tau_{21}\ \delta x_1\ \delta x_2\ \delta x_3$ in the negative x_3 direction.

When we look at the tractions acting on the right and left faces, as shown in Figure 1.8c, the couple that results (see Fig. 1.8d) is $\tau_{12}\ \delta x_1\ \delta x_2\ \delta x_3$ in the positive x_3 direction. No other couple is acting in the x_3 direction, so the two couples we have obtained must be equal and opposite, otherwise the cube would spin up with increasing angular velocity. It follows that $\tau_{21} = \tau_{12}$.

By similar arguments requiring no net couple about the x_1 or x_2 directions, we find $\tau_{32} = \tau_{23}$ and $\tau_{31} = \tau_{13}$. So in general, $\tau_{ji} = \tau_{ij}$ and we have proven the symmetry required to obtain (1.17) from (1.16).

1.1.2 NORMAL STRESS AND SHEAR STRESS

Figure 1.9 shows two components of a traction vector $\mathbf{T}$, one in the normal direction, and the other in the direction parallel to the surface across which $\mathbf{T}$ acts. The normal stress, σ_n, is given by

$$\begin{aligned}\sigma_n &= \text{component of traction in the normal direction}\\ &= \mathbf{T}\cdot\mathbf{n}\\ &= T_j n_j\\ &= \tau_{ij} n_i n_j.\end{aligned} \tag{1.18}$$

Note that normal stress, as defined here, is a scalar. But it is associated with a particular direction, and the vector $\sigma_n \mathbf{n}$ is often used for the normal stress (as shown in Figure 1.9).

The shear stress is the component of $\mathbf{T}$ in the plane of the surface across which traction is acting, so it may be evaluated by subtracting the normal stress $\sigma_n \mathbf{n}$ from $\mathbf{T}$ itself:

$$\text{shear stress} = \mathbf{T} - (\mathbf{T}\cdot\mathbf{n})\ \mathbf{n}. \tag{1.19}$$

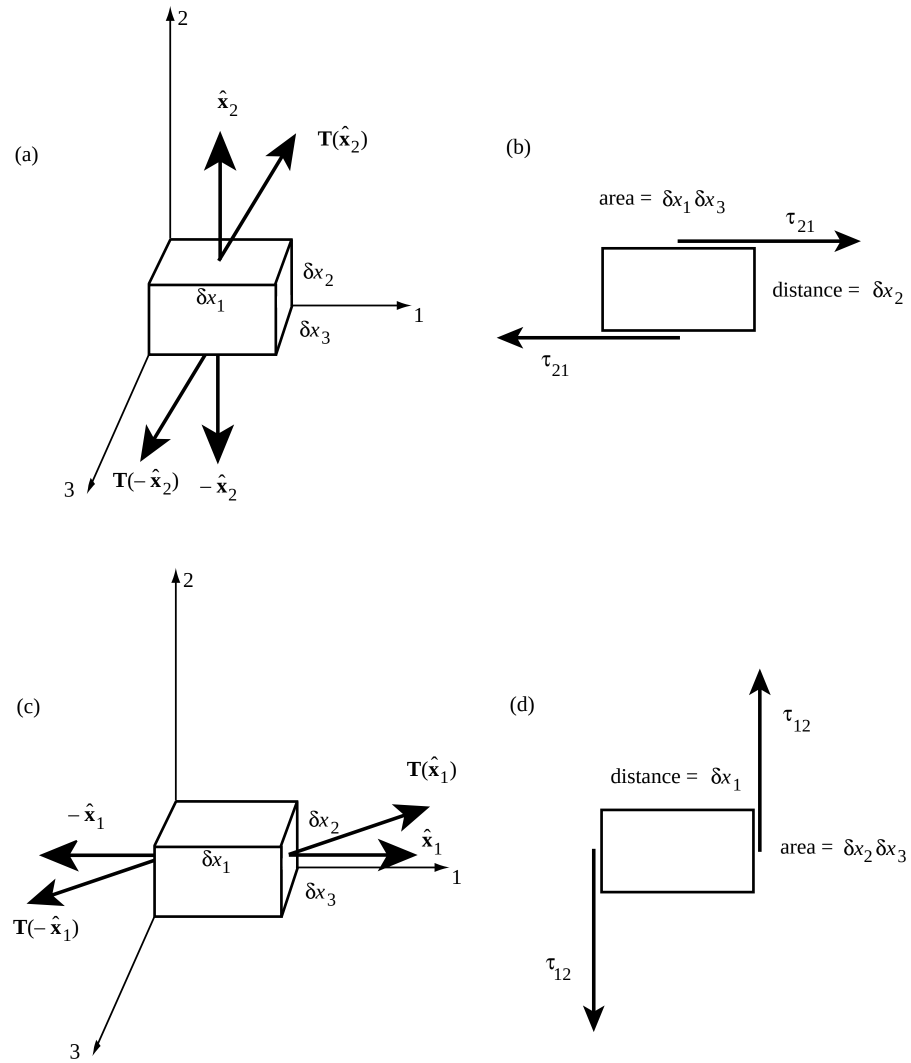

FIGURE **1.8**
A small cube is shown, with an analysis of the couple tending to make the cube spin about the Ox_3 axis. Thus, in (a) is shown the tractions acting on the top and bottom faces. In (b) is shown the couple associated with these tractions which will tend to make the cube rotate about the axis into the plane of the page — the couple consists of a pair of forces $\tau_{21}\delta x_1\ \delta x_3$ separated by a distance δx_2. In (c) is shown the tractions acting on the right and left faces, and in (d) the associated couple tending to make the cube rotate about an axis out of the plane of the page.

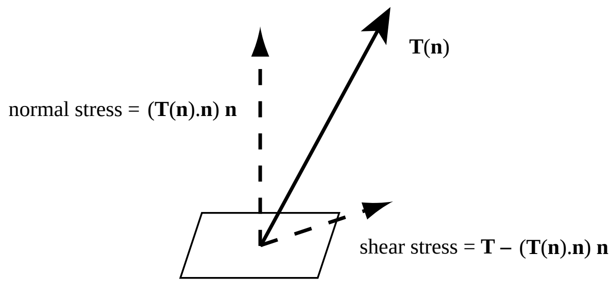

FIGURE **1.9**
The traction **T** is shown resolved into its normal and shear components.

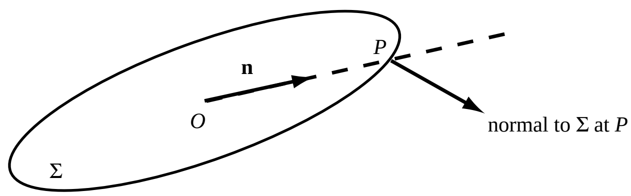

FIGURE **1.10**
Σ is a "quadric surface" centered on the coordinate origin O. P is the point on Σ at the place where a line drawn from the origin in the direction of **n** (a unit vector) meets the quadric surface. The normal to Σ at P is shown, and in general it will be in a direction different from the direction of **n**.

It is unfortunate that the word "stress" is used for scalar quantities as in (1.18), for vector quantities as in (1.19), and for tensor quantities as in (1.15).

In a fluid with no viscosity, the scalar normal stress is equal to pressure (but with opposite sign, in our convention where traction is positive if it is pulling). And in such a fluid, the shear stresses are all zero.

1.1.3 A QUADRIC SURFACE ASSOCIATED WITH THE STRESS TENSOR

The surface Σ given by the equation

$$\tau_{ij} x_i x_j = \text{constant} \tag{1.20}$$

has properties that are independent of the coordinate system used to define components x_i and τ_{ij}.

Σ is a three-dimensional surface, either a spheroid or a hyperboloid (which may have one or two separate surfaces). Because (1.20) has no linear terms (terms in x_i only), Σ is a surface which is symmetric about the origin O, as shown in Figure 1.10. (If **x** is a point on the surface Σ, then so is the point $-\mathbf{x}$.)

If P is at position **x**, then

BOX **1.1**

On finding the direction of a normal to a surface

If Σ is the surface composed of points $\mathbf{x}$ which satisfy the equation

$$F(\mathbf{x}) = c$$

for some constant c, then consider a point $\mathbf{x} + \delta\mathbf{x}$ which also lies on Σ.

Then $F(\mathbf{x} + \delta\mathbf{x}) = c$, so

$$F(\mathbf{x} + \delta\mathbf{x}) - F(\mathbf{x}) = 0.$$

But

$$\begin{aligned} F(\mathbf{x} + \delta\mathbf{x}) &= F(\mathbf{x}) + \delta x_i \frac{\partial F}{\partial x_i} \qquad \text{(Taylor series expansion)} \\ &= F(\mathbf{x}) + \delta\mathbf{x} \cdot \nabla F. \end{aligned}$$

Therefore,

$$\delta\mathbf{x} \cdot \nabla F = 0$$

for all directions $\delta\mathbf{x}$ such that $\mathbf{x}$ and $\mathbf{x} + \delta\mathbf{x}$ lie in the surface Σ. Since $\delta\mathbf{x}$ is tangent to Σ at the point $\mathbf{x}$, and $\delta\mathbf{x} \cdot \nabla F = 0$ for all such tangents, this last equantion shows that ∇F has the defining property of the normal to Σ at $\mathbf{x}$ (it is perpendicular to all the tangents at $\mathbf{x}$).

We conclude that the normal to the surface $F(\mathbf{x}) = c$ is parallel to ∇F.

$$\text{vector } OP = \mathbf{x} = OP\ \mathbf{n} \tag{1.21}$$

and the length OP is related to the normal stress. This follows, because

$$\begin{aligned} \text{constant} &= \tau_{ij} x_i x_j \\ &= \tau_{ij} OP\ n_i OP\ n_j \\ &= OP^2\ \tau_{ij} n_i n_j \\ &= OP^2\ \sigma_n \qquad \text{from (1.18).} \end{aligned}$$

Hence,

$$\sigma_n \propto \frac{1}{OP^2}. \tag{1.22}$$

This is a geometrical property of Σ, which depends only the shape of the quadric surface, not its absolute size.

Another such geometrical property is associated with the normal to Σ at P. We can use Box 1.1 to see that this normal is in the direction whose i-th component is

$$\begin{aligned}
\frac{\partial}{\partial x_i}(\tau_{kl}x_k x_l) &= \tau_{kl}\frac{\partial x_k}{\partial x_i}x_l + \tau_{kl}x_k\frac{\partial x_l}{\partial x_i} \qquad \text{(chain rule)} \\
&= \tau_{kl}\delta_{ik}x_l + \tau_{kl}x_k\delta_{il} \quad \text{(the partial derivatives are Kronecker delta functions)} \\
&= \tau_{il}x_l + \tau_{ki}x_k \qquad \text{(using the substitution property (1.7))} \\
&= 2\tau_{ij}x_j \quad \text{(using symmetry of the stress tensor, and changing dummy subscripts)} \\
&= 2OP\ \tau_{ij}n_j \qquad \text{(from (1.21))} \\
&= 2OP\ T_i \qquad \text{(from (1.17))}.
\end{aligned}$$

It follows that

$$\text{the normal to } \Sigma \text{ at } P \text{ is parallel to the traction vector, } \mathbf{T}. \tag{1.23}$$

The results (1.22) and (1.23) are the two key geometrical features of the quadric surface representing the underlying tensor $\boldsymbol{\tau}$, and they are independent of any coordinate system. Going back to Figures 1.4 and 1.9, the quadric surface shown in Figure 1.10 is a geometrical device for obtaining the direction of the traction vector, and the way in which normal stress varies as the unit normal **n** varies (specifying the orientation of an area element across which the traction acts).

We can for example see geometrically that there are three special directions **n**, for which the traction $\mathbf{T}(\mathbf{x})$ is parallel to **n** and hence (for these **n** directions) the shear stress vanishes. To obtain the same result algebraically, we note that these special values of **n** are such that $T_i \propto n_i$ and so, using (1.17),

$$\tau_{ij}n_j = \lambda n_i. \tag{1.24}$$

As discussed in Box 1.2, this last equation has solutions (for **n**), but only if λ takes on one of three special values (eigenvalues); and the three resulting values of **n** (one for each λ value) are mutually orthogonal.

1.1.4 FORMAL DEFINITION OF A SECOND ORDER CARTESIAN TENSOR

If two cartesian coordinate systems $Ox_1x_2x_3$ and $Ox_1'x_2'x_3'$ are related to each other as shown in Figure 1.3, with direction cosines defined by $l_{ij} = \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j$, then the entity **A** is a second order cartesian tensor if its components A_{ik} in the $Ox_1x_2x_3$ system and A'_{jl} in the $Ox_1'x_2'x_3'$ system are related to each other by

$$A'_{jl} = A_{ik}l_{ij}l_{kl}. \tag{1.25}$$

It is left as an exercise (in Problem 1.1) to obtain the reverse transformation, and to show that the associated quadric surface $A_{ij}x_ix_j = c$ is given also by $A'_{ij}x'_ix'_j = c$.

The tensor **A** is *isotropic* if its components are the same in the original and the transformed coordinate systems. It follows geometrically that the quadric surface of such a tensor is simply a sphere, and that $A_{ij} = A'_{ij} = \text{constant} \times \delta_{ij}$. The stress tensor in an inviscid fluid is isotropic (Problem 1.4 asks you to prove this).

BOX **1.2**
On solutions of

$$\mathbf{A} \cdot \mathbf{x} = \lambda \mathbf{x}$$

where **A** *is a symmetric real matrix*

Here we briefly review the main properties of the set of linear equations

$$A_{ij} x_j = \lambda x_i \qquad (\text{where } A_{ji} = A_{ij};\ i = 1, \ldots, n; \text{ and } j = 1, \ldots, n). \tag{1}$$

These are n equations in n unknowns. In the present chapter, usually $n = 3$.

Equation (1) is fundamentally different from the standard linear problem $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, or

$$A_{ij} x_i = b_i, \tag{2}$$

which has a straightforward solution provided $\det \mathbf{A} \neq 0$, for then the inverse matrix exists and $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$. Note that for equation (1) there is no restriction on the absolute size of $\mathbf{x}$: if $\mathbf{x}$ is a solution of (1), then so is $c\mathbf{x}$ for any constant c. But if $\mathbf{x}$ is a solution of (2), then in general $c\mathbf{x}$ will *not* be a solution.

Equation (1) has the trivial solution $\mathbf{x} = \mathbf{0}$, but we are interested in non-trivial solutions, for which at least one component of $\mathbf{x}$ is non-zero. It follows that equation (1) represents a set of n scalar relations between $n - 1$ variables, for example x_j / x_1 $(j = 2, \ldots, n)$. If $x_1 = 0$ then we can instead divide (1) by some other component of $\mathbf{x}$.

We can use the Kronecker delta function to write (1) as

$$(A_{ij} - \lambda \delta_{ij}) x_j = 0, \qquad \text{or as} \qquad (\mathbf{A} - \lambda \mathbf{I}) \cdot \mathbf{x} = \mathbf{0}, \tag{1, again}$$

where $\mathbf{I}$ is the identity matrix, $I_{ij} = \delta_{ij}$, with 1 at every entry on the diagonal and 0 everywhere else.

If $\det(\mathbf{A} - \lambda \mathbf{I}) \neq 0$, then the inverse $(\mathbf{A} - \lambda \mathbf{I})^{-1}$ exists and the only solution of (1) is $\mathbf{x} = (\mathbf{A} - \lambda \mathbf{I})^{-1} \cdot \mathbf{0}$, giving $\mathbf{x} = \mathbf{0}$. It follows that the only way we can obtain non-trivial solutions of (1), is to require that

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \tag{3}$$

Equation (3) is an n-th order polynomial in λ, and in general it has n solutions, $\lambda^{(\alpha)}$ $(\alpha = 1, \ldots, n)$, called *eigenvalues* of the matrix $\mathbf{A}$. For each eigenvalue, we have an associated *eigenvector* $\mathbf{x}^{(\alpha)}$ satisfying

$$\mathbf{A} \cdot \mathbf{x}^{(\alpha)} = \lambda^{(\alpha)} \mathbf{x}^{(\alpha)} \qquad (\text{not summed over } \alpha). \tag{4}$$

These eigenvectors (there are n of them, since $\alpha = 1, \ldots, n$) are the vector solutions of (1), so they have the special property that $\mathbf{A} \cdot \mathbf{x}^{(\alpha)}$ is in the same direction as the vector $\mathbf{x}^{(\alpha)}$ itself. Equation (1) imposes no contraint on the absolute size of solutions such as $\mathbf{x}^{(\alpha)}$, and we are free to normalize these solutions if we wish. A common choice is to make $\mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\alpha)} = 1$.

The final important property of the eigenvectors, is that any two eigenvectors corresponding to two different eigenvalues must be orthogonal. To prove this, we consider two eigenvalues $\lambda^{(\alpha)}$ and $\lambda^{(\beta)}$ with corresponding eigenvectors $\mathbf{x}^{(\alpha)}$ and $\mathbf{x}^{(\beta)}$, and $\lambda^{(\alpha)} \neq \lambda^{(\beta)}$. Then

$$A_{ij} x_j^{(\alpha)} = \lambda^{(\alpha)} \mathbf{x}_i^{(\alpha)} \qquad \text{and} \qquad A_{ij} x_j^{(\beta)} = \lambda^{(\beta)} \mathbf{x}_i^{(\beta)}. \tag{5}$$

BOX **1.2** (continued)

We can multiply both sides of the first of equations (5) by $x_i^{(\beta)}$ and sum over i, and multiply both sides of the second of equations (5) by $x_i^{(\alpha)}$ and sum over i, giving

$$A_{ij}x_j^{(\alpha)}x_i^{(\beta)} = \lambda^{(\alpha)}x_i^{(\alpha)}x_i^{(\beta)} = \lambda^{(\alpha)}\mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\beta)} \tag{6a}$$

and

$$A_{ij}x_j^{(\beta)}x_i^{(\alpha)} = \lambda^{(\beta)}x_i^{(\alpha)}x_i^{(\beta)} = \lambda^{(\beta)}\mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\beta)}. \tag{6b}$$

Note here that we are using the Einstein summation convention in the usual way for repeated subscripts, but we are *not* using it for repeated superscripts α and β. [The summation convention does not work naturally with such superscripts, because one side of equations such as (4) or (6a) has repeated superscripts, and the other side has the superscript in just one place.]

Because **A** is a symmetric matrix, $A_{ij}x_j^{(\beta)}x_i^{(\alpha)} = A_{ji}x_j^{(\beta)}x_i^{(\alpha)}$. But we can use any symbol for repeated subscripts, and in particular we can exchange the symbols i and j in this last expression, so that $A_{ij}x_j^{(\beta)}x_i^{(\alpha)} = A_{ji}x_j^{(\beta)}x_i^{(\alpha)} = A_{ij}x_j^{(\beta)}x_i^{(\alpha)} = A_{ij}x_i^{(\alpha)}x_j^{(\beta)}$. Hence, the left-hand side of (6a) equals the left-hand side of (6b). Subtracting (6b) from (6a), we see that

$$(\lambda^{(\alpha)} - \lambda^{(\beta)})\, \mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\beta)} = 0.$$

But the first of these factors is not zero, since we took the eigenvalues to be different ($\lambda^{(\alpha)} \neq \lambda^{(\beta)}$). It follows that the second factor must be zero, $\mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\beta)} = 0$, and hence that the eigenvectors corresponding to different eigenvalues are orthogonal.

The best way to become familiar with the eigenvalue/eigenvector properties described above, is to work through the details of a few examples (such as Problem 1.2).

The above formal definition of a tensor is rarely useful directly, as a way to demonstrate that an entity (suspected of being a tensor) is in fact a tensor. For this purpose, we usually rely instead upon the so-called *quotient rule*, which states that if the relationship $T_{ik}A_k = B_i$ is true is all coordinate systems, where A_k and B_i are the components of vectors **A** and **B**, then the T_{ik} are components of a second order cartesian tensor provided the components of **A** can all be varied separately. The statement that "the relationship $T_{ik}A_k = B_i$ is true is all coordinate systems" means that $T'_{jk}A'_l = B'_j$ as well as $T_{ik}A_k = B_i$.

To show that the quotient rule means the formal definition is satisfied, note that

$$\begin{aligned} T'_{jk}A'_l &= B'_j \\ &= B_i l_{ij} && \text{since } \mathbf{B} \text{ is a vector} \\ &= T_{ik}A_k l_{ij} && \text{using } T_{ik}A_k = B_i \\ &= T_{ik}A'_l l_{kl} l_{ij} && \text{from the reverse transform } A_k = A'_l l_{kl} \text{ — see Problem 1.1.} \end{aligned}$$

Hence

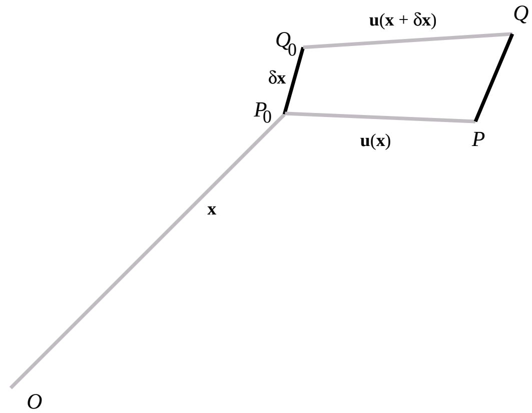

FIGURE **1.11**
A line element $\delta\mathbf{x}$ is shown with ends at P_0 and Q_0, and also after the neighborhood of P_0 has undergone deformation. The new position of the line element is from P to Q. P_0 is at $\mathbf{x}$, Q_0 is at $\mathbf{x} + \delta\mathbf{x}$, P is at $\mathbf{x} + \mathbf{u}$ and Q is at $\mathbf{x} + \delta\mathbf{x} + \mathbf{u}(\mathbf{x} + \delta\mathbf{x})$.

$$(T'_{jl} - T_{ik} l_{ij} l_{kl}) A'_l = 0. \tag{1.26}$$

This is a set of three scalar equations $(j = 1, 2, \text{or } 3)$. But since they are also true as the components A'_l are varied independently, it follow that the coefficients of the $A'_l (l = 1, 2, \text{ or } 3)$ must all vanish. So $T'_{jl} = T_{ik} l_{ij} l_{kl}$, and $\mathbf{T}$ satisfies the formal definition of a tensor (compare with (1.25)).

The entity we have been calling the stress tensor, $\boldsymbol{\tau}$, satisfies the quotient rule (see equation (1.17), in which the components of $\boldsymbol{\tau}$ relate the traction vector and $\mathbf{n}$). It is for this reason that indeed we are justified in referring to $\boldsymbol{\tau}$ as a second order cartesian tensor.

1.2 The strain tensor

Two different methods are widely used to describe the motions and the mechanics of motion in a continuum. These are the Lagrangian description, which emphasizes the study of a particular particle that is specified by its original position at some reference time, and the Eulerian description, which emphasizes the study of whatever particle happens to occupy a particular spatial location. Note that a seismogram is the record of motion of a particular part of the Earth (namely, the particles to which the seismometer was attached during installation), so it is directly a record of Lagrangian motion. A pressure gauge attached to the sea floor also provides a Lagrangian record, as does a neutrally buoyant gauge that is free to move in the water. But a gauge that is fixed to the sea floor and measuring properties (such as velocity, temperature, opacity) of the water flowing by, provides an Eulerian record.

We use the term *displacement*, regarded as a function of space and time, and written as $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$, to denote the vector distance of a particle at time t from the position $\mathbf{x}$ that it occupies at some reference time t_0, often taken as $t = 0$. Since $\mathbf{x}$ does not change with time, it follows that the *particle velocity* is $\partial\mathbf{u}/\partial t$ and that the *particle acceleration* is $\partial^2\mathbf{u}/\partial t^2$.

To analyze the distortion of a medium, whether it be solid or fluid, elastic or inelastic, we use the *strain tensor*. If a particle initially at position $\mathbf{x}$ is moved to position $\mathbf{x} + \mathbf{u}$, then the relation $\mathbf{u} = \mathbf{u}(\mathbf{x})$ is used to describe the displacement field. To examine the distortion of the part of the medium that was initially in the vicinity of $\mathbf{x}$, we need to know the new position of the particle that was initially at $\mathbf{x} + \delta\mathbf{x}$, where $\delta\mathbf{x}$ is a small line-element. This new position (see Figure 1.11) is $\mathbf{x} + \delta\mathbf{x} + \mathbf{u}(\mathbf{x} + \delta\mathbf{x})$. Any distortion is liable to change the relative position of the ends of the line-element $\delta\mathbf{x}$. If this change is $\delta\mathbf{u}$, then $\delta\mathbf{x} + \delta\mathbf{u}$ is the new vector line-element, and by writing down the difference between its end points we obtain

$$\delta\mathbf{x} + \delta\mathbf{u} = \mathbf{x} + \delta\mathbf{x} + \mathbf{u}(\mathbf{x} + \delta\mathbf{x}) - (\mathbf{x} + \mathbf{u}(\mathbf{x})).$$

Since $|\delta\mathbf{x}|$ is arbitrarily small, we can expand $\mathbf{u}(\mathbf{x} + \delta\mathbf{x})$ as $\mathbf{u} + (\delta\mathbf{x} \cdot \nabla)\mathbf{u}$ plus negligible terms of order $|\delta\mathbf{x}|^2$. It follows that $\delta\mathbf{u}$ is related to gradients of $\mathbf{u}$ and to the original line-element $\delta\mathbf{x}$ via

$$\delta\mathbf{u} = (\delta\mathbf{x} \cdot \nabla)\mathbf{u}, \qquad \text{or} \qquad \delta u_i = \frac{\partial u_i}{\partial x_j}\,\delta x_j. \tag{1.27}$$

However, we do not need all of the nine independent components of $\frac{\partial u_i}{\partial x_j}$ to specify true distortion in the vicinity of $\mathbf{x}$, since part of the motion is due merely to an infinitesimal rigid-body rotation of the neighborhood of $\mathbf{x}$. We shall write

$$\delta u_i = \frac{\partial u_i}{\partial x_j}\delta x_j = \tfrac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\delta x_j + \tfrac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)\delta x_j$$

and use the result given in Problem 1.8 to interpret the last term of the above equation. Introducing the notation $u_{i,j}$ for $\frac{\partial u_i}{\partial x_j}$, we then see that (1.27) can be rewritten as

$$\delta u_i = \tfrac{1}{2}(u_{i,j} + u_{j,i})\delta x_j + \tfrac{1}{2}(\text{curl}\,\mathbf{u} \times \delta\mathbf{x})_i, \tag{1.28}$$

and the rigid-body rotation is of amount $\frac{1}{2}\text{curl}\ \mathbf{u}$. The interpretation of the last term in (1.28) as a rigid-body rotation is valid if $\left|u_{i,j}\right| \ll 1$. If displacement gradients were not "infinitesimal" in the sense of this inequality, then we would instead have to analyze the contribution to $\delta\mathbf{u}$ from a *finite* rotation—a much more difficult matter, since finite rotations do not commute and cannot be expressed as vectors.

In terms of the infinitesimal strain tensor, defined to have components

$$e_{ij} \equiv \tfrac{1}{2}(u_{i,j} + u_{j,i}), \tag{1.29}$$

the effect of true distortion on any line-element $\delta\mathbf{x}$ is to change the relative position of its end points by a displacement whose i-th component is $e_{ij}\delta x_j$. By the quotient rule discussed in Section 1.1.4, it follows that $\mathbf{e}$ is indeed a second order cartesian tensor. Rotation does not affect the length of the element, and the new length is

$$\begin{aligned}|\delta\mathbf{x}+\delta\mathbf{u}| &= \sqrt{\delta\mathbf{x}\cdot\delta\mathbf{x}+2\delta\mathbf{u}\cdot\delta\mathbf{x}} && \text{(neglecting } \delta\mathbf{u}\cdot\delta\mathbf{u}\text{)}\\ &= \sqrt{\delta x_i\delta x_i + 2e_{ij}\delta x_i\delta x_j} && \text{(from (1.28), and using (curl } \mathbf{u}\times\delta\mathbf{x})\cdot\delta\mathbf{x}=0\text{)}\\ &= |\delta\mathbf{x}|\,(1+e_{ij}\nu_i\nu_j) && \text{(to first order, if } \left| e_{ij} \right| \ll 1\text{),}\end{aligned}$$

where $\boldsymbol{\nu}$ is the unit vector $\delta\mathbf{x}/\,|\delta\mathbf{x}|$. It follows that the extensional strain of a line-element originally in the $\boldsymbol{\nu}$ direction, which we define to be

$$e(\boldsymbol{\nu}) = \frac{\text{change in length}}{\text{original length}} = \frac{PQ - P_0Q_0}{P_0Q_0},$$

is given by

$$e(\boldsymbol{\nu}) = e_{ij}\nu_i\nu_j. \tag{1.30}$$

Strain is a dimensionless quantity. The diurnal solid Earth tide leads to strains of about 10^{-7}. Strains of about 10^{-11} due to seismic surface waves from distant small earthquakes can routinely be measured with modern instruments.

1.2.1 THE STRAIN QUADRIC

We can define the surface $e_{ij}x_ix_j =$ constant. If it is referred to the axes of symmetry as coordinate axes, this quadric surface becomes $E_1x_1^2 + E_2x_2^2 + E_3x_3^2 =$ constant and the strain tensor components become

$$\mathbf{e} = \begin{pmatrix} E_1 & 0 & 0\\ 0 & E_2 & 0\\ 0 & 0 & E_3\end{pmatrix}.$$

The $E_i (i = 1, 2, \text{ or } 3)$, called *principal strains*, are eigenvalues of the matrix

$$\mathbf{e} = \begin{pmatrix} e_{11} & e_{12} & 3_{13}\\ e_{12} & e_{22} & 3_{23}\\ e_{13} & e_{23} & 3_{33}\end{pmatrix}$$

(this matrix is different for different cartesian coordinate systems, but its eigenvalues are the same). Figure 1.12 shows an example of the strain quadric, in the case that the E_i do not all have the same sign. In this case, Σ is a hyperboloid. The figure caption describes the two key geometrical properties of Σ, analogous to the results shown previously for the stress quadric (see (1.22) and (1.23)). The first property, $e(\nu) \propto OP^{-2}$, follows from taking $\mathbf{x} = OP\boldsymbol{\nu}$ and substituting for the components of $\mathbf{x}$ in $e_{ij}x_ix_j =$ constant. The second property, concerning the direction of the displacement due to distortion, follows from showing that the direction of the normal at P is parallel to $\nabla(e_{ij}x_ix_j)$, which has i-th component $2e_{ij}x_j = 2OPe_{ij}\nu_j$. In terms of the line-element $\delta\mathbf{x}$, this i-th component is proportional to $e_{ij}\delta x_j$, and indeed this is in the displacement direction due to the distortion (see comments following (1.29)).

The physical interpretation of the axes of symmetry of Σ, is that these are the three special (mutually orthogonal) directions in which an original line-element is merely

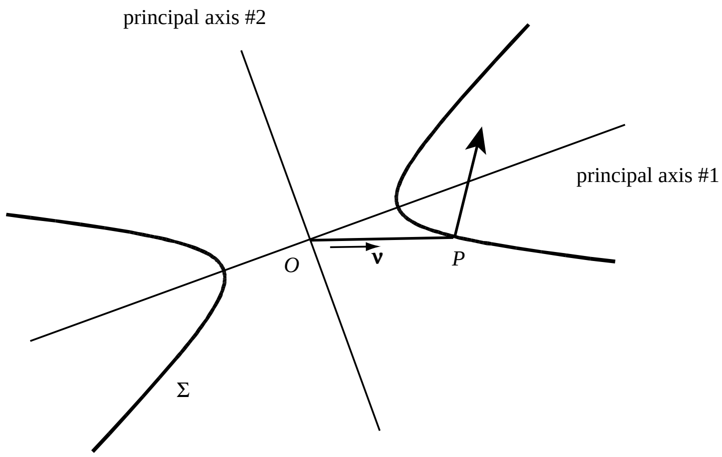

FIGURE **1.12**
A hyperboloid is shown here, as an example of a strain quadric Σ. The third principal axis is perpendicular to the first two (out of the page). For a line-element in the ν direction, extensional strain is inversely proportional to OP^2. The displacement of the end of the line-element, due to distortion, is in the direction of the normal at P.

shortened or lengthened by the deformation. For these special directions, line-elements are not subject to any shearing motions. The only rotation, is that which applies to the whole neighborhood of the line-element as a rigid body rotation (distinct from deformation).

1.3 Some simple examples of stress and strain

Examples are

(i) Compacting sediments, which shrink in the vertical direction but stay the same in the horizontal direction. Think of them as being in a large tank (see Figure 1.13). If x_1 and x_2 are horizontal directions with x_3 vertical, then the strain tensor has components

$$\mathbf{e} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & e_{33} \end{pmatrix}.$$

Compression at depth, from the weight of those above, leads to $e_{33} < 0$.

But although the strain tensor has only one non-zero component, the stress tensor has non-zero values of τ_{11}, τ_{22}, τ_{33}. If the sediments lack strength, then the stress will essentially be the same as if the material were composed of a fluid (lacking any ability to resist shearing forces). In this case, the stress at depth x_3 is due solely to pressure generated by the overburden. There are no shearing forces, and

$$\boldsymbol{\tau} = -\begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} \qquad \text{where } P = \int_0^{x_3} \rho g \, dx_3.$$

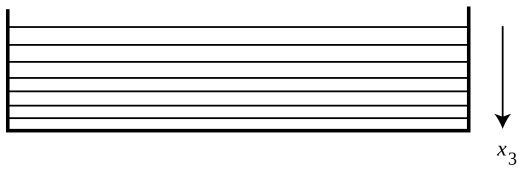

FIGURE **1.13**
Sediments, compressed vertically.

(ii) A wire that is stretched in the x_1−direction will shrink in the perpendicular directions x_2 and x_3. The strain tensor is

$$\mathbf{e} = \begin{pmatrix} e_{11} & 0 & 0 \\ 0 & e_{22} & 0 \\ 0 & 0 & e_{33} \end{pmatrix}, \quad \text{with } e_{11} > 0 \text{ and } e_{22} = e_{33} < 0.$$

The stress tensor has only one non-zero component:

$$\boldsymbol{\tau} = \begin{pmatrix} \tau_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(iii) We use the term *pure shear* to describe the type of deformation shown in Figure 1.14, in which a small rectangle is subjected to the stress field

$$\boldsymbol{\tau} = \begin{pmatrix} 0 & \tau_{12} & 0 \\ \tau_{12} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In this deformation, the point B has moved to the right from its original position, by an amount $\dfrac{\partial u_1}{\partial x_2}\delta x_2$, and the point A has moved up an amount $\dfrac{\partial u_2}{\partial x_1}\delta x_1$. In the case of pure shear, there is no rigid body rotation. So curl $\mathbf{u} = \mathbf{0}$ which means here that $\dfrac{\partial u_1}{\partial x_2} = \dfrac{\partial u_2}{\partial x_1}$. The line element OB_0 has been rotated over to the right (clockwise) by an angle given by $\dfrac{\partial u_1}{\partial x_2}\delta x_2 \div \delta x_2 = \dfrac{\partial u_1}{\partial x_2}$. The line element OA_0 has been rotated anticlockwise by the angle $\dfrac{\partial u_2}{\partial x_1}\delta x_1 \div \delta x_1 = \dfrac{\partial u_2}{\partial x_1} = \dfrac{\partial u_1}{\partial x_2}$ since there is no rigid body rotation, and the strain tensor components are

$$\mathbf{e} = \begin{pmatrix} 0 & e_{12} & 0 \\ e_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Always, $e_{ji} = e_{ij}$. And for pure shear, $e_{12} = e_{21} = \dfrac{\partial u_1}{\partial x_2} = \dfrac{\partial u_2}{\partial x_1}$.

The reduction in the original $\pi/2$ angle at O (Figure 1.14) is $2e_{12}$. The ratio between stress and this angle reduction,

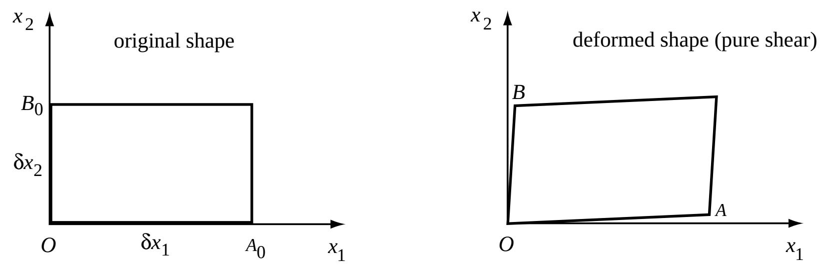

FIGURE **1.14**
A small rectangular prism is sheared with no change in area. The shearing stresses τ_{12} and $\tau_{21} (= \tau_{12})$ are the same as those shown in Figure 1.8b and 1.8d.

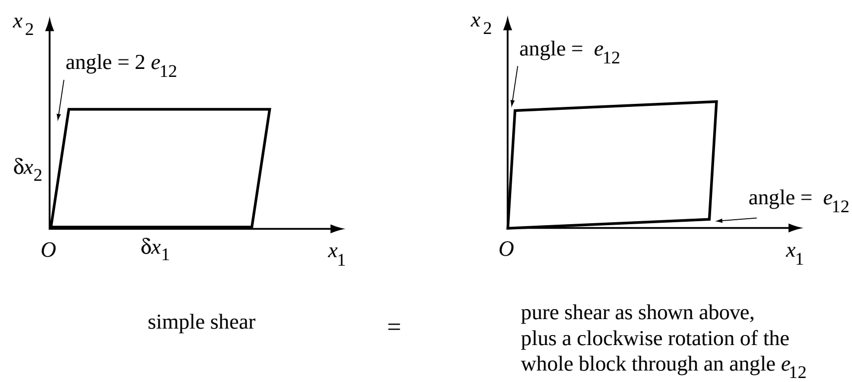

FIGURE **1.15**
A simple shear is shown on the left, with $u_1 \propto x_2$, so that an original rectangle in the Ox_1x_2-plane is now sheared with motion only in the x_1-direction. This motion is equivalent to a pure shear followed by a clockwise rigid-body rotation. The angle reduction at the origin is $2e_{12}$ in both cases. Rigidity is the ratio between shear stress and the angle reduction. For a viscous material, the material continues to deform as long as the shear stress is applied, and viscosity is defined as the ratio between shear stress and the *rate* of angle reduction.

$$\mu = \frac{\tau_{12}}{2e_{12}}, \tag{1.31}$$

is known as the *rigidity*.

(iv) Consider a deformation in which the displacement is $\mathbf{u} = (u_1, 0, 0)$ and u_1 depends on $\mathbf{x}$ only via the x_2-component. An example is shown in Figure 1.15, and this situation is called a *simple shear*. It can be regarded as a pure shear described by $e_{12} = e_{21} = \frac{1}{2}\frac{\partial u_1}{\partial x_2}$, plus a rigid rotation of amount $\frac{1}{2}$ curl $\mathbf{u} = \left(0, 0, -\frac{1}{2}\frac{\partial u_1}{\partial x_2}\right)$.

The definition of rigidity is easier to understand with a simple shearing deformation. It is the shearing stress τ_{21} divided by the displacement gradient $\frac{\partial u_1}{\partial x_2}$ (equal to the angle reduction as shown in Figure 1.15).

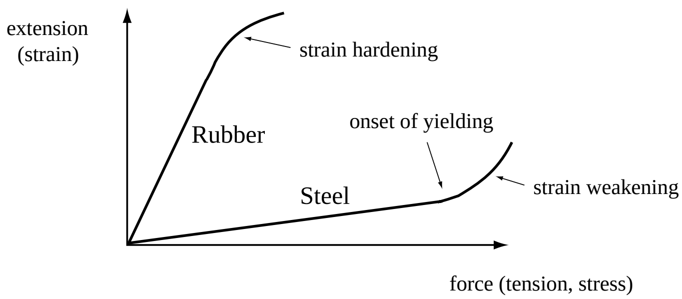

FIGURE **1.16**
For some solid materials, strain grows in proportion to stress at low stress, but the linearity is lost and strain either increases more slowly with increasing stress (for example, rubber), or increases more quickly (for example, steel). At high enough stress, the material breaks. This is a schematic diagram: some solids break at extensional strains as small as 10^{-3}, and some forms of rubber maintain linearity even for strains of order 1.

All of the above examples in this subsection are *elastic* examples. Stress and strain go to zero together, if $\tau_{ij} = c_{ijkl}e_{kl}$, and there is no time-dependence.

Real materials can be quite different, exhibiting viscous behavior, or strain hardening, or a tendency to yield at high values of applied stress (see Figure 1.16). We take up this subject next, with emphasis on solids.

1.4 Relations between stress and strain

A medium is said to be *elastic* if it possesses a natural state (in which strains and stresses are zero) to which it will revert when applied forces are removed. Under the influence of applied loads, stress and strain will change together, and the relation between them, called the *constitutive relation*, is an important characteristic of the medium. Over 300 years ago, Robert Hooke summarized the relationship today known as Hooke's Law between the extension of a spring, and the force acting to cause the extension. He concluded experimentally that force $\propto$ extension. The constant of proportionality here is often called a *modulus*, M (say), and then

$$\text{force} = M \times \text{extension}.$$

(Hooke's Law appeared originally as an anagram, *ceiiinosssttuv*, of the Latin phrase *ut tensio, sic vis* — meaning "as the extension, so the force." Some scientific personalities are very strange.)

The modern generalization of Hooke's law is that each component of the stress tensor is a linear combination of all components of the strain tensor, i.e., that there exist constants c_{ijkl} such that

$$\tau_{ij} = c_{ijpq}e_{pq}. \tag{1.32}$$

A material that obeys the constitutive relation (1.32) is said to be *linearly elastic*. The quantities c_{ijkl} are components of a fourth-order tensor, and have the symmetries

$$c_{jipq} = c_{ijpq} \qquad (\text{due to } \tau_{ji} = \tau_{ij}) \tag{1.33}$$

and

$$c_{ijqp} = c_{ijpq} \qquad (\text{due to } e_{qp} = e_{pq}). \tag{1.34}$$

It is also true from a thermodynamic argument that

$$c_{pqij} = c_{ijpq}, \tag{1.35}$$

for a material in which the energy of deformation (associated with tensors $\boldsymbol{\tau}$ and $\mathbf{e}$) does not depend on the time history of how the deformation was acquired.

The c_{ijkl} are independent of strain, which is why they are sometimes called "elastic constants," although they are varying functions of position in the Earth. In general, the symmetries (1.33), (1.34), and (1.35) reduce the number of independent components in c_{ijkl} from 81 to 21. There is considerable simplification in the case of an isotropic medium, since **c** must be isotropic. It can be shown that the most general isotropic fourth-order tensor, having the symmetries of **c**, has the form

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{1.36}$$

This involves only two independent constants, λ and μ, known as the Lamé moduli. Substituting from (1.36) into (1.32), we see that the stress–strain relation becomes

$$\tau_{ij} = \lambda\delta_{ij}e_{kk} + 2\mu e_{ij} \tag{1.37}$$

in isotropic elastic media.

If we consider only shearing stresses and shearing strains, using (1.37) with $i \neq j$, then $\tau_{ij} = 2\mu e_{ij}$ and μ here is just the rigidity we introduced earlier, in (1.31). For many materials, the Lamé moduli λ and μ (sometimes called Lamé constants) are approximately equal. They can be used to generate a number of other constants that characterize the properties of material which are subjected to particular types of strain and stress, for example the stretched wire discussed as item (ii) of Section 1.3. For that example, (1.37) gives

$$\begin{aligned} \tau_{11} &= \lambda(e_{11} + e_{22} + e_{33}) + 2\mu e_{11} \\ 0 &= \lambda(e_{11} + e_{22} + e_{33}) + 2\mu e_{22} \\ 0 &= \lambda(e_{11} + e_{22} + e_{33}) + 2\mu e_{33}. \end{aligned} \tag{1.38}$$

By comparing the second and third of these equations, we see that $e_{22} = e_{33}$; and then from these same two equations it follows that $\lambda e_{11} + 2(\lambda + \mu)e_{22} = 0$. *Poisson's ratio* is defined as

$$\nu = \frac{\text{shrinking strain}}{\text{stretching strain}}$$

in this example of simple stretching, so

$$\nu = \frac{-e_{22}}{e_{11}} = \frac{\lambda}{2(\lambda + \mu)} \sim \frac{1}{4} \tag{1.39}$$

if λ and μ are approximately equal (which is the case for many common materials).

Another important class of materials is that for which shearing stresses lead to flow. Such materials are not elastic. We introduce the *deviatoric stress* as the difference between τ_{ij} and the average of the principal stresses, which is $\frac{1}{3}\tau_{kk}$ and which we can symbolize by $-P$. (In an inviscid or perfect fluid, P with this definition is simply the scalar pressure.) The *viscosity* of a material that can flow, such as syrup (with low viscosity) or the Earth's mantle (with high viscosity, and significant flow occurring only over many millions of years), is the constant of proportionality between the deviatoric stress and twice the strain rate. Thus, the viscosity ν is given by

$$\tau_{ij} + P\delta_{ij} = 2\nu\dot{e}_{ij}. \tag{1.40}$$

A factor 2 appears here, for the same reason that a factor 2 appears in the definition of rigidity (see (1.31)): both rigidity and viscosity are better appreciated for simple shearing motions, than for pure shearing. For a viscous material, simple shearing due to application of τ_{21} as shown in Figure 1.15 results in angle reduction of an original rectangle at the rate $\frac{\partial \dot{u}_1}{\partial x_2} = 2\dot{e}_{12}$. Viscosity is shear stress divided by the spatial gradient of particle velocity.

Suggestions for Further Reading

Menke, William, and Dallas Abbott. *Geophysical Theory*, pp 41–50 for basic properties of tensors, pp 237–252 for properties of strain and stress tensors. New York: Columbia University Press, 1990.

Problems

1.1 Consider the two sets of cartesian axes shown in Figure 1.3, with l_{ij} as the direction cosine of the angle between Ox_i and Ox'_j. Thus, the unit vector $\mathbf{x}'_1$ expressed in $Ox_1x_2x_3$ coordinates has the components (l_{11}, l_{21}, l_{31}).

a) Show in general that $l_{ij}l_{ik} = \delta_{jk}$ and also that $l_{ij}l_{kj} = \delta_{ik}$.
b) Hence show that $\left(\mathbf{l}^{-1}\right)_{ij} = l_{ji}$.
c) We know from (1.4) that a vector $\mathbf{V}$ having components V_i in the $Ox_1x_2x_3$ system has components in the $Ox'_1x'_2x'_3$ system given by $V'_j = l_{ij}V_i$. Show that the original components are given in terms of the transformed components by

$$V_k = l_{kj}V'_j.$$

d) For two vectors $\mathbf{a}$ and $\mathbf{b}$ with components defined in both of the two coordinate systems, show that $a_ib_i = a'_jb'_j$.

e) We know from Section 1.1.4 that components of **A** in the two coordinate systems are related by $A'_{jl} = A_{ik} l_{ij} l_{kl}$. Show that $A'_{jl} l_{pj} l_{ql} = A_{pq}$ and hence that the reverse transformation can be written as $A_{ik} = A'_{jl} l_{ij} l_{kl}$.

f) The quadric surface associated with **A** is $A_{ij} x_i x_j = c$ in the $Ox_1x_2x_3$ system. Show that this same surface is $A'_{ij} x'_i x'_j = c$ in the $Ox'_1x'_2x'_3$ system.

1.2 Find the eigenvalues and eigenvectors of the following matrices:

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 0.11 & 0.48 \\ 0.48 & 0.39 \end{pmatrix},$$

and

$$\mathbf{A} = \begin{pmatrix} 144p + 25q & -60(p-q) \\ -60(p-q) & 25p + 144q \end{pmatrix}$$

(normalize the eigenvectors, so that $\mathbf{x}^{(\alpha)} \cdot \mathbf{x}^{(\alpha)} = 1$).

1.3 Consider a stress tensor $\boldsymbol{\tau}$ whose components in a particular cartesian coordinate system are

$$\boldsymbol{\tau} = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

a) What are the principal stresses $\sigma_1, \sigma_2, \sigma_3$ in this coordinate system (order them so that $\sigma_1 < \sigma_2 < \sigma_3$)? And what are the corresponding three mutually orthogonal unit vector directions $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$, $\boldsymbol{\nu}_3$?

b) Evaluate the system of direction cosines

$$\begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

where l_{ij} is the cosine of the angle between $\hat{\mathbf{x}}_i$ and ν_j (this matrix is not symmetric).

c) Show that in the present example the matrix of l_{ij} values has the property described in Problem 1.1ab, namely that pre-multiplication or post-multiplication of **l** by its transpose gives the identity matrix, and hence that $\left(\mathbf{l}^{-1}\right)_{ij} = l_{ji}$.

d) Show in this example that when the principal axes are taken as coordinates, the components of $\boldsymbol{\tau}$ become

$$\boldsymbol{\tau} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}.$$

[Note then that if the principal axes are taken as the coordinates $OX_1X_2X_3$, the equation of the stress quadric is simply $\sigma_1 X_1^2 + \sigma_2 X_2^2 + \sigma_3 X_3^2 = \text{constant}$.]

1.4 Suppose that at a point inside a fluid with volume V the stress field has the property that the normal stress across any surface through the point is a constant, $-P$

(independent of the orientation of the surface). Show then that the corresponding stress tensor at a point in a fluid is isotropic, and has components $\tau_{ij} = -P\delta_{ij}$.

1.5 Show that

$$e_{kk} = u_{l,l} = \nabla \cdot \mathbf{u} = \frac{\text{increase in volume}}{\text{original volume}}$$

for a small volume element of material that is deformed. [*Hint*: to prove the result, consider a small cube with faces parallel to the principal axes of strain (the symmetry axes of the strain tensor). Alternatively, use the physical definition of divergence (flux out of a volume, per unit volume). The result itself is the reason that e_{kk} and $\nabla \cdot \mathbf{u}$ are sometimes referred to as "volumetric strain."]

1.6 The *bulk modulus* κ of an isotropic material is defined as

$$\kappa = \frac{-P}{\text{volumetric strain}}$$

when the material is compressed by an all-round pressure P.

a) Show that $\kappa = \lambda + \frac{2}{3}\mu$ where λ and μ are the Lamé moduli.
b) κ is sometimes given other names, such the *compressibility* or the *incompressibility*. Which of these two names makes more sense?

1.7 Use the alternating tensor ε_{ijk} discussed in (1.6) and (1.7) to show for vectors **a**, **b**, and **c**, that

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}.$$

Using $(\nabla^2 \mathbf{u})_i = \nabla^2(u_i)$, show also that

$$\nabla^2 \mathbf{u} = \nabla(\nabla \cdot \mathbf{u}) - \nabla \times (\nabla \times \mathbf{u}).$$

[This last result, obtained here by using cartesian components, is essentially a definition of $\nabla^2 \mathbf{u}$, that can be applied to non-cartesian components.]

1.8 Using $u_{i,j}$ to denote $\dfrac{\partial u_i}{\partial u_j}$, show that $\varepsilon_{ijk}\varepsilon_{jkm}u_{m,l}\delta x_k = (u_{i,j} - u_{j,i})\delta x_j$ and hence that $(u_{i,j} - u_{j,i})\delta x_j = (\text{curl}\,\mathbf{u} \times \delta\mathbf{x})_i$.

1.9 Show that Poisson's ratio (ν) is $\frac{1}{2}$ for a material that is incompressible.

1.10 Young's modulus (E) is defined as the ratio of stretching stress to stretching strain in the example (ii) of Section 1.3 (see also (1.39)). Show that

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}.$$

1.11 For an isotropic elastic material, the stress is given in terms of strains by (1.37). Show that strain is then given in terms of the stresses by

$$2\mu e_{ij} = -\frac{\lambda}{3\lambda + 2\mu}\tau_{kk}\delta_{ij} + \tau_{ij}.$$