MATH 323 Cartesian Tensors Module

Chapter 1 – Definition and properties of Cartesian Tensors

Contents:

Revision

Newton's laws of motion

Newton's three laws are:

1. A body at rest or in uniform motion in a straight line will continue to be so unless acted on by an external force.
2. The acceleration a of the body is proportional to the force F and inversely proportional to its mass m. That is:

   a = F/m   or   F = m a
3. To every action (force) there is an equal and opposite reaction. If you press down on the floor with a force equal to your weight, then the floor presses up on you with the same force.

Vectors, projections, scalar and vector products

A vector is an ordered n-tuple of numbers $(v_1, v_2, \ldots v_n) \in R^n$. We write this $\underline{\mathbf{v}}$. We can display $\underline{\mathbf{v}}$ as a row or a column – generally we shall prefer to deal with *column* vectors, so we write:

$$\underline{\mathbf{v}} = (v_1, v_2, \ldots v_n)^T.$$

where T denotes transpose.

A special subset of vectors are position vectors $\underline{\mathbf{x}} = (x_1, x_2, x_3)^T \in R^3$, which gives the Cartesian coordinates of a point in (non-relativistic) Euclidean space.

We can equally write $\underline{x}$ as $\underline{\mathbf{x}} = (x_1, x_2, x_3)^T$

or $\underline{\mathbf{x}} = x_1 \mathbf{u_1} + x_2 \mathbf{u_2} + x_3 \mathbf{u_3}$

where { $\mathbf{u_1}\ \mathbf{u_2}\ \mathbf{u_3}$ } is the set of *unit* vectors that point along the coordinate axes; so that e.g $\mathbf{u_2} = (0, 1, 0)^T$.

The *scalar product* of two vectors $\underline{\mathbf{w}}$, $\underline{\mathbf{v}}$ is defined to be

$$\underline{\mathbf{w}} \bullet \underline{\mathbf{v}} = \sum_j w_j v_j = \underline{\mathbf{w}}^T \underline{\mathbf{v}} = \underline{\mathbf{v}}^T \underline{\mathbf{w}}$$

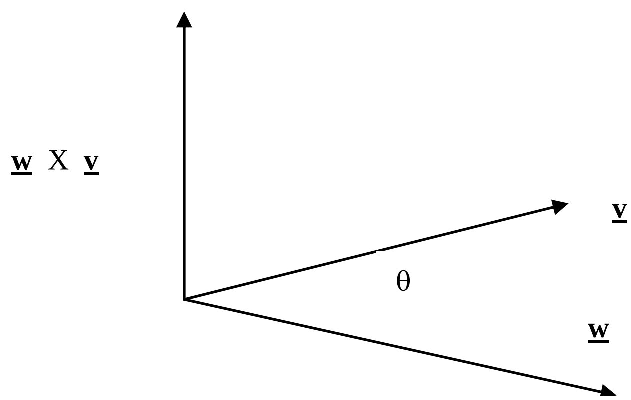
NB this is numerically equally to $|\underline{\mathbf{w}}|\,|\underline{\mathbf{v}}| \cos\theta$, where θ is the angle between the vectors.

The *vector product* of two vectors $\underline{\mathbf{w}}, \underline{\mathbf{v}} \in R^3$ is defined to be a vector whose components may be calculated using

$$\underline{\mathbf{w}} \times \underline{\mathbf{v}} = \begin{vmatrix} \mathbf{u_1} & \mathbf{u_2} & \mathbf{u_3} \\ \mathbf{w_1} & \mathbf{w_2} & \mathbf{w_3} \\ \mathbf{v_1} & \mathbf{v_2} & \mathbf{v_3} \end{vmatrix}$$

where | | means 'determinant'. E.g. the 2nd component of $\underline{\mathbf{w}} \times \underline{\mathbf{v}} = \mathbf{w_3 v_1 - w_1 v_3}$.

The *magnitude* of this vector is given by $|\underline{\mathbf{w}}|\,|\underline{\mathbf{v}}| \sin\theta$, (θ is the angle between the vectors) and its direction is a vector at right angles to both $\underline{\mathbf{w}}$ and $\underline{\mathbf{v}}$ whose direction is given by the *right-hand rule*.

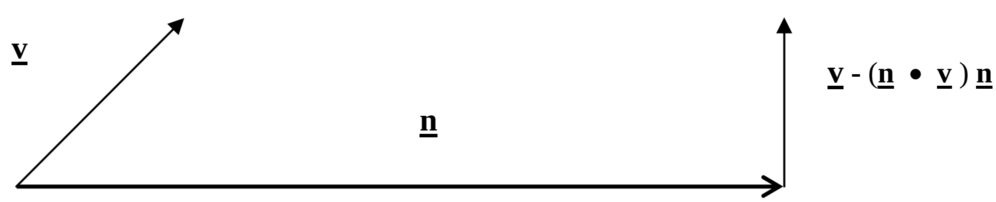


We will define 'direction' to *mean* the unit vector (often $\underline{\mathbf{n}}$) that points in the direction we want.

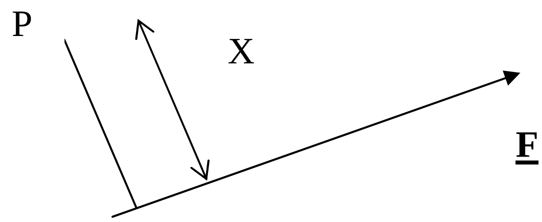
The projection of a vector ***on*** any direction is found by taking the scalar product of the vector with the direction.

The vector that is the projection of $\underline{\mathbf{v}}$ ***in the direction*** $\underline{\mathbf{n}}$ is $\underline{\mathbf{v}} - (\underline{\mathbf{n}} \bullet \underline{\mathbf{v}})\,\underline{\mathbf{n}}$.

E.g. the projection of $\underline{\mathbf{v}} = (v_1, v_2, v_3)$ along the x_2 axis is $\underline{\mathbf{v}} - [(0, 1, 0) \bullet \underline{\mathbf{v}}]\,(0, 1, 0) = (v_1, 0, v_3)$.



Moment of a force, or ***torque***, about a point P = the product F X of the force **F** and the perpendicular distance X of the force to the point. Viz.



Newton's 2nd Law for rotational motion

$$\tau = I \alpha$$

where $\boldsymbol{\tau}$ is the torque vector, $\boldsymbol{\alpha}$ is the angular acceleration and I is the 3x3 Inertia matrix.

Taylor Series (multivariate version)

If $f(\mathbf{x})$ is a function of $\mathbf{x} \in R^n$ that is differentiable m times at a point $\mathbf{x_0}$, then

$$f(\mathbf{x}) = f(\mathbf{x_0}) + \sum_j \partial f(\mathbf{x_0})/\partial x_j \, (x_j - x_{0j}) + 1/2! \sum_j \sum_k \partial^2 f(\mathbf{x_0})/\partial x_j \partial x_k \, (x_j - x_{0j})(x_k - x_{0k})$$

plus higher terms and a remainder term $R^m(\mathbf{x})$.

Eigenvectors and eigenvalues

Let A be a real square matrix. $A \mathbf{x} = \lambda \mathbf{x}$ has a non-trivial solution if and only if $| A - \lambda I | = 0$, where $|\;\;|$ means determinant. The solutions λ and $\mathbf{x}$ are called the eigenvalues and eigenvectors of A.

Exercise: 1. Prove this theorem (easy).

2. Find the eigenvalues λ and corresponding eigenvectors $\mathbf{x}$ for the matrix:

$$\begin{bmatrix} 2 & 0 & 6 \\ 0 & 1 & 0 \\ 6 & 0 & -4 \end{bmatrix}$$

===

Indexed sets

We have met vectors $\underline{\mathbf{v}} = (v_1, v_2, v_3, \ldots\ldots v_n)^T \in R^n$, and matrices:

$A = \{ A_{ij}, i = 1,\ldots n, j = 1,\ldots n\};$

We want to generalise this to an arbitrary number of indices, eg:

$A = \{ a_{ijk\ldots p}, i = 1,\ldots n, j = 1,\ldots n, \ldots. p = 1,\ldots. n\}.$

We will associate an indexed set A with a typical element of it $a_{ijk\ldots p}$, so that $a_{ijk\ldots p}$ means *either* an element of A *or* the whole of A. Eg for a vector we write v_i to mean either an element of $\underline{\mathbf{v}}$ or the whole of $\underline{\mathbf{v}}$. The context will make clear which we mean.

Summation convention

We will be dealing with equations linking indexed sets. A familiar example involves the multiplication of a vector $\underline{\mathbf{x}}$ by a matrix A, where we can write the components of the product:

$$y_i = \sum_{j=1}^{N} A_{ij} x_j \qquad (1)$$

where, of course, the number of columns, N, of A must equal the dimension of $\underline{\mathbf{x}}$.

The summation convention (ref. A. Einstein) says that in an equation involving index sets (and later, tensors) we can omit the summation symbol, and write in place of 1:

$$y_i = A_{ij} x_j \qquad (1a)$$

There are some rules associated with this convention:

(1) In an index set equation, a repeated index implies summation over the dimension of set (usually 3 in our case). This repeated index is therefore a *dummy index*, and may be *replaced by any convenient symbol*.

(2) An index may not appear three or more times on the same side of an index set equation.
E.g. $y_i = A_{ij} x_j z_j$ is meaningless, because we are unsure what is intended to be summed.

(3) Dummy indices aside, the same set of (true) indices must appear on both sides of an index set equation.
E.g. $b_{ik} = A_{ij} x_j z_k$ is valid;
$b_{ik} = A_{ij} z_k$ is not.

Exercise: which of the following are valid index set expressions?

(i) $a_{ij} = b_{ijk} c_{jk}$

(ii) $a_{iji} = b_{jk} c_k$

(iii) $a = b_{ij} c_{jk} b_{ik}$

(iv) $a_i d_{ii} = b_{ij} c_j$

Fundamental principle of description of physical laws and properties

It is a fundamental principle of science that a physical property – such as gravity, magnetic field, or state of strain – and laws describing behaviour and interactions of physical quantities, must be independent of the system of spatial coordinates we use to measure it in or relate it to. For example, the gravitational field at a point on the Earth is the same whether we use Cartesian coordinates or latitude and longitude to describe the location of the point.

So an indexed set represents a physical property if and only if this property is unchanged when we change to a different coordinate system.

The components of the vector may change as we change coordinate systems, but the vector's fundamental properties, magnitude and direction, will not.

This leads us to consider changes of coordinate system and to ask *which indexed sets remain unchanged* in the sense above when the coordinate system changes.

Cartesian Coordinates

We will mostly be dealing with Cartesian coordinate systems. However, the concepts are transferable to any coordinate system e.g. spherical polar coordinates.

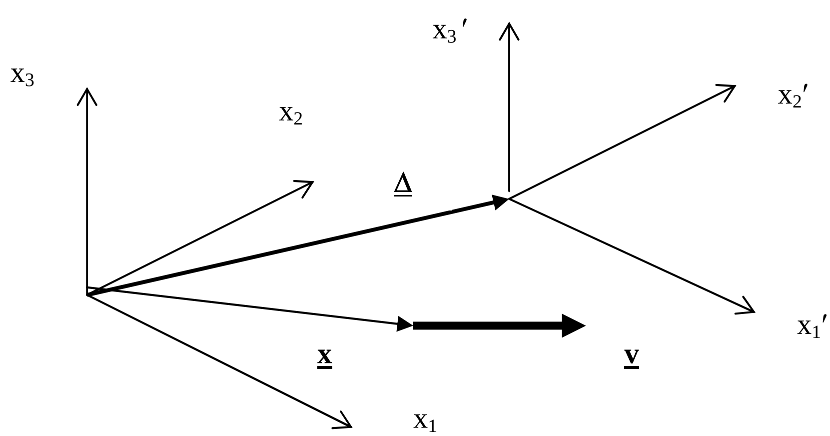
What changes in Cartesian coordinates are possible?

We rule out changes of scale, as this is merely a change of measuring units (eg metres to feet). We rule out distortions of the coordinate axes – we insist on rectilinear Cartesian coordinates with axes at right angles. This leaves three possible changes of axes:

- Translation - change of origin without re-orientation of the axes;
- Rotation of the axes about the origin;
- Reflection in some plane;

or any combination of these.

We dismiss translation as being trivial. We can illustrate this by considering a vector field $\underline{\mathbf{v}} = v_i(\underline{\mathbf{x}})$ defined at any point $\underline{\mathbf{x}}$ in R^3.



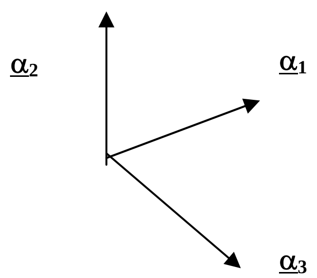
Note that we represent the coordinate axes by x_1, x_2, x_3, rather than x, y, z, and the components of any (column) vector by $(v_1, v_2, v_3)^T$, where T is transpose. This enables us to identify $\underline{\mathbf{v}}$ with v_i.

If the origin of coordinates shifts by $\underline{\Delta}$, the point $\underline{\mathbf{x}}$ becomes $\underline{\mathbf{x}} - \underline{\Delta}$ in the new (′) coordinate system, the point $\underline{\mathbf{x}} + \underline{\mathbf{v}}$ becomes $\underline{\mathbf{x}} + \underline{\mathbf{v}} - \underline{\Delta}$, but $\underline{\mathbf{v}} = (\underline{\mathbf{x}} + \underline{\mathbf{v}} - \underline{\Delta}) - (\underline{\mathbf{x}} - \underline{\Delta})$ is unchanged.

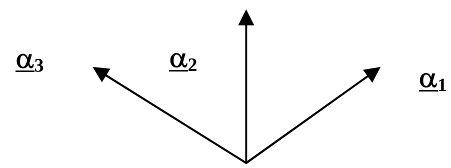
This then leaves us with rotations and reflections to consider.

Representations of coordinate axes

We can represent a general set of three mutually perpendicular axes by a set of three mutually perpendicular *unit length* vectors $\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$.



or



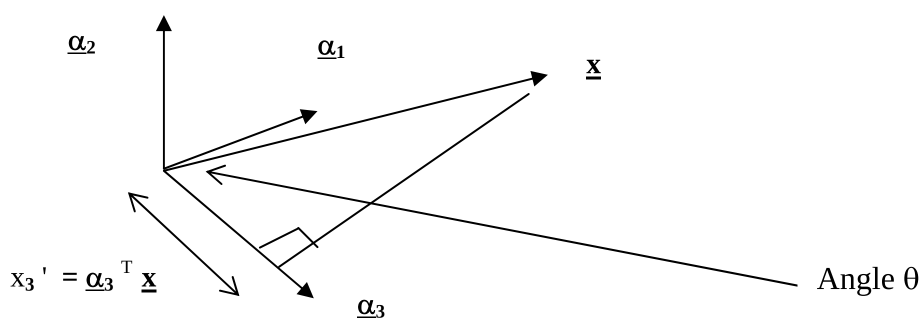
(Right Hand Set) (Left Hand Set)

E.g. the ‘standard’ Cartesian coordinate axes are represented by the vectors $(1, 0, 0)^T$, $(0, 1, 0)^T$ and $(0, 0, 1)^T$.

A Right Handed set of coordinates is one where $\underline{\alpha}_i \times \underline{\alpha}_j = \underline{\alpha}_k$, where i, j, k is any cyclic permutation of 1, 2, 3 and **x** is cross product. If the axes do not obey this rule, they are a Left Handed set.

Now we want to consider how the description of any position $\underline{\mathbf{x}} = (x_1 , x_2 , x_3)^T$ changes when we change from the standard coordinate system to one represented by $\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$. In the new coordinate system, the components of $\underline{\mathbf{x}}$ can be found by projecting $\underline{\mathbf{x}}$ onto the new coordinate direction, i.e. taking scalar products of $\underline{\mathbf{x}}$ with each of the coordinate vectors: $\underline{\alpha}_1^T \underline{\mathbf{x}}$, $\underline{\alpha}_2^T \underline{\mathbf{x}}$, $\underline{\alpha}_3^T \underline{\mathbf{x}}$.

E.g.



Since $|\underline{\alpha}_i| = 1$, $\underline{\alpha}_3^T \underline{\mathbf{x}} = \underline{\alpha}_3 \bullet \underline{\mathbf{x}} = |\underline{\mathbf{x}}| \cos\theta = x_3'$, the 3rd component of $\underline{\mathbf{x}}$ in the ‘new’ coordinate system ($\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$). Similarly for the 1st and 2nd components.

So we have the three components of the position of $\underline{\mathbf{x}}$ in the new system i.e. we can write

$$\underline{\mathbf{x}}' = \begin{bmatrix} \underline{\alpha_1}^T \underline{\mathbf{x}} \\ \underline{\alpha_2}^T \underline{\mathbf{x}} \\ \underline{\alpha_3}^T \underline{\mathbf{x}} \end{bmatrix}$$

where the prime ′ denotes 'new coordinates'.

Write

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

And think of each column of A as a vector, viz:

$$A = (\underline{\alpha_1}, \quad \underline{\alpha_2}, \quad \underline{\alpha_3})$$

ie $\underline{\alpha_i} = \begin{bmatrix} A_{1i} \\ A_{2i} \\ A_{3i} \end{bmatrix}$

Then

$$A^T \underline{\mathbf{x}} = \begin{bmatrix} \underline{\alpha_1}^T \\ \underline{\alpha_2}^T \\ \underline{\alpha_3}^T \end{bmatrix} (\underline{\mathbf{x}})$$

$$= \begin{bmatrix} \underline{\alpha_1}^T \underline{\mathbf{x}} \\ \underline{\alpha_2}^T \underline{\mathbf{x}} \\ \underline{\alpha_3}^T \underline{\mathbf{x}} \end{bmatrix} = \underline{\mathbf{x}}'$$

So we can write the three components of $\underline{\mathbf{x}}$ in the new coordinate system by:

$$\underline{\mathbf{x}}' = A^T \underline{\mathbf{x}} \qquad (2)$$

Now consider

$$A^T A = \begin{vmatrix} \underline{\alpha_1}^T \\ \underline{\alpha_2}^T \\ \underline{\alpha_3}^T \end{vmatrix} (\underline{\alpha_1}, \quad \underline{\alpha_2}, \quad \underline{\alpha_3})$$

$$= \begin{bmatrix} \underline{\alpha_1}^T \underline{\alpha_1}, & \underline{\alpha_1}^T \underline{\alpha_2}, & \underline{\alpha_1}^T \underline{\alpha_3} \\ \underline{\alpha_2}^T \underline{\alpha_1}, & \underline{\alpha_2}^T \underline{\alpha_2}, & \underline{\alpha_2}^T \underline{\alpha_3} \\ \underline{\alpha_3}^T \underline{\alpha_1}, & \underline{\alpha_3}^T \underline{\alpha_2}, & \underline{\alpha_3}^T \underline{\alpha_3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad = \qquad I$$

because the $\underline{\alpha_i}$ are mutually orthogonal unit vectors i.e. $\underline{\alpha_i}^T \underline{\alpha_j} = 0$ if $i \neq j$,
or $\underline{\alpha_i}^T \underline{\alpha_j} = 1$ if $i = j$.

We say that a linear transformation

$$\underline{\mathbf{x}}' = A^T \underline{\mathbf{x}} ,$$

with the property that $A^T A = I$, is *orthogonal.*

Orthogonal transformations and their properties

A^T and A^{-1}

An implication of $A^T A = I$ is that $A^T = A^{-1}$, i.e.

$$A^{-1} A = I$$

But this implies $A A^{-1} = I$, so $A A^T = I$ also. Note that this means that the *rows* of A are mutually orthogonal unit vectors, as well as the columns. We will discuss their interpretation below.

Determinant of A

For $A = (\underline{\alpha_1}, \underline{\alpha_2}, \underline{\alpha_3})$, the determinant of A, |A|, equals the vector triple product:

$$|A| = \underline{\alpha_1} \bullet \underline{\alpha_2} \times \underline{\alpha_3}$$

But for a RH set $\underline{\alpha_2} \times \underline{\alpha_3} = \underline{\alpha_1}$, so $|A| = \underline{\alpha_1} \bullet \underline{\alpha_1} = 1$ (or –1 if $(\underline{\alpha_1}, \underline{\alpha_2}, \underline{\alpha_3})$ were a LH set).

Reverse transformation

From eqn (2),

$$\underline{\mathbf{x}} = A \, \underline{\mathbf{x}}' \qquad (2a)$$

which describes the transformation from the new coordinates $\underline{\mathbf{x}}'$ back to the standard ones.

Successive transformations and transformations between arbitrary Cartesian coordinate systems.

Let $(\underline{\alpha_1}, \underline{\alpha_2}, \underline{\alpha_3})$, $(\underline{\beta_1}, \underline{\beta_2}, \underline{\beta_3})$ be two sets of mutually orthogonal unit vectors representing two sets of coordinate axes. If $\underline{\mathbf{x}}$ is the position of a point in the $\underline{\alpha}$ coordinates, what is it in the $\underline{\beta}$?

We calculate this by going via the standard coordinates ($\underline{\mathbf{x}}_1$, $\underline{\mathbf{x}}_2$, $\underline{\mathbf{x}}_3$). As in (2a), the coordinates of $\underline{\mathbf{x}}$ in standard coordinates are:

$$\underline{\mathbf{x}}_s \quad = \quad A\ \underline{\mathbf{x}}$$

Let

$$B \quad = \quad (\underline{\beta}_1 , \underline{\beta}_2 , \underline{\beta}_3)$$

Then from (2),

$$\underline{\mathbf{x}}' \quad = \quad B^T \underline{\mathbf{x}}_s$$

gives the coordinates of $\underline{\mathbf{x}}_s$ in the ($\underline{\beta}_1$, $\underline{\beta}_2$, $\underline{\beta}_3$) system. Therefore

$$\underline{\mathbf{x}}' \quad = \quad B^T A\ \underline{\mathbf{x}}$$

gives the transformation that takes the description of $\underline{\mathbf{x}}$ from the ($\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$) system to the ($\underline{\beta}_1$, $\underline{\beta}_2$, $\underline{\beta}_3$) system.

Note that $(B^T A)^T B^T A = A^T B B^T A = A^T I A = I$; i.e. $B^T A$ is an orthogonal transformation.

Interpreting the rows and columns of an orthogonal transformation

Put $C = B^T A$. What is the significance of the rows and columns of C?
Repeating:

$$\underline{\mathbf{x}}' \quad = \quad C\ \underline{\mathbf{x}}$$

gives the transformation that takes the description of $\underline{\mathbf{x}}$ from the ($\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$) system to the ($\underline{\beta}_1$, $\underline{\beta}_2$, $\underline{\beta}_3$) system.

Let $\underline{\mathbf{x}}$ be one of the coordinate vectors $\underline{\alpha}_i$; say $\underline{\alpha}_1$. In the ($\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$) system, therefore,

$$\underline{\mathbf{x}} = (1, 0, 0)^T$$

Therefore

$$C\ \underline{\mathbf{x}} = C \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} C_{11} \\ C_{21} \\ C_{31} \end{vmatrix}$$

i.e. the first column of C. But this is the description of $(1, 0, 0)^T$ in the new, $\underline{\beta}$, coordinate system as described by transformation C. That is, the columns of C are the descriptions of the original coordinate axes in the new, $\underline{\beta}$ system.

Similarly

$$\underline{\mathbf{x}} \quad = \quad C^T \underline{\mathbf{x}}'$$

Now let $\underline{\mathbf{x}}'$ be one of the coordinate vectors $\underline{\beta}_i$; say $\underline{\beta}_2$. Therefore,

$$\underline{\mathbf{x}}' = (0, 1, 0)^T$$

Therefore

$$C^T \underline{\mathbf{x}}' = C^T \begin{vmatrix} 0 \\ 1 \\ 0 \end{vmatrix} = \begin{vmatrix} C_{21} \\ C_{22} \\ C_{23} \end{vmatrix}$$

i.e. the second row of C. But this is the description of $(0, 1, 0)^T$ in the original, $\underline{\alpha}$, coordinate system as described by transformation C. That is the rows of C are the descriptions of the new coordinate axes in the original system system.

Reflections

We have already seen that an orthogonal transformation $A = (\underline{\alpha}_1 , \underline{\alpha}_2 , \underline{\alpha}_3)$, $A^T A = A A^T = I$, represents a new coordinate system whose axes lie along the directions of the (unit) vectors $\underline{\alpha}_i$, as described in the old coordinate system.

If the new system represented a change from, say, a RH system to a left-handed one e.g. $\underline{\alpha}_1 \times \underline{\alpha}_2 = - \underline{\alpha}_3$, then

$$A^* = (\underline{\alpha}_1 , \underline{\alpha}_2 , - \underline{\alpha}_3)$$

is also an orthogonal transformation which preserves the right-handedness. We can transform from the first new system to the second one by reflecting in the $\underline{\alpha}_1$, $\underline{\alpha}_2$ plane, using the orthogonal transformation:

$$A_{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

ie $A A_{(2)} = A^*$.

All reflections can be represented in this way. If we wish to reflect in a plane P that does not correspond to the plane of two of the existing coordinate axes, then we can first rotate the coordinates to make the plane P coincide with the plane of a pair of new coordinates and then reflect.

This means that we can, in practice, ignore reflections. We are thus left with rotations of the coordinate axes as *the only non-trivial change of coordinate system we have to consider*.

Euler's Theorem for rigid bodies with one point fixed

To illustrate some of these ideas, we will prove a famous theorem of Euler's that has special application in Earth science.

Consider two *distinct* sets of mutually orthogonal vectors:

$$A = (\underline{\alpha}_1 , \underline{\alpha}_2 , \underline{\alpha}_3) , \quad B = (\underline{\beta}_1 , \underline{\beta}_2 , \underline{\beta}_3)$$

(i.e. we do not allow $\underline{\alpha}_i = \underline{\beta}_i$, for all i).

Any vector $\underline{\mathbf{x}}$ (described with respect to the standard Cartesian axes) has its components in the $(\underline{\alpha}_1 , \underline{\alpha}_2 , \underline{\alpha}_3)$ and $(\underline{\beta}_1 , \underline{\beta}_2 , \underline{\beta}_3)$ coordinate systems given by

$$\underline{\mathbf{x}}_\alpha = A^T \underline{\mathbf{x}} ; \quad \underline{\mathbf{x}}_\beta = B^T \underline{\mathbf{x}}$$

as we have seen.

Can we find a (non-trivial) vector $\underline{\mathbf{x}}$ that has the same components in the two 'new' systems?
i.e. find $\underline{\mathbf{x}}$ such that $\underline{\mathbf{x}}_{\alpha} = \underline{\mathbf{x}}_{\beta}$?

For such an x,

$$A^T \underline{\mathbf{x}} = B^T \underline{\mathbf{x}}$$

or $$A^T \underline{\mathbf{x}} - B^T \underline{\mathbf{x}} = (A^T - B^T)\underline{\mathbf{x}} = 0 \qquad (3)$$

which is satisfied if $\underline{\mathbf{x}} = 0$ (trivial case) or if

$$(\underline{\alpha}_i - \underline{\beta}_i)^T \underline{\mathbf{x}} = 0 \qquad i = 1, 2, 3 \qquad (4)$$

There is a difficulty with equation 4. At first sight it appears to give three equations (i = 1, 2, 3) for the three unknown components of $\underline{\mathbf{x}}$. However, if $\underline{\mathbf{x}}$ satisfies eqn (3) so does $k\,\underline{\mathbf{x}}$, i.e. $\underline{\mathbf{x}}$ is only a direction, with *two* independent components. So equation (4) may *overdetermine* $\underline{\mathbf{x}}$, which would lead to the conclusion that (3) has only the trivial solution.

The first two parts of eqn (4) are;

$$(\underline{\alpha}_1 - \underline{\beta}_1)^T \underline{\mathbf{x}} = 0 \quad \text{and} \quad (\underline{\alpha}_2 - \underline{\beta}_2)^T \underline{\mathbf{x}} = 0$$

The third part is:

$$(\underline{\alpha}_3 - \underline{\beta}_3)^T \underline{\mathbf{x}} = 0 \qquad (4 - 3)$$

The first two parts say that $\underline{\mathbf{x}}$ is orthogonal to both $(\underline{\alpha}_1 - \underline{\beta}_1)$ and $(\underline{\alpha}_2 - \underline{\beta}_2)$.

This means that $\underline{\mathbf{x}}$ is parallel to $(\underline{\alpha}_1 - \underline{\beta}_1)\,\mathbf{X}\,(\underline{\alpha}_2 - \underline{\beta}_2)$ (because the two sets are distinct this cross product cannot be zero).

so write $\underline{\mathbf{x}} = k\,(\underline{\alpha}_1 - \underline{\beta}_1)\,\mathbf{X}\,(\underline{\alpha}_2 - \underline{\beta}_2)$.

Now since $(\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3)$ and $(\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3)$ are RH sets of unit vectors,

$$\underline{\alpha}_1 \,\mathbf{X}\, \underline{\alpha}_2 = \underline{\alpha}_3 \quad \text{and} \quad \underline{\beta}_1 \,\mathbf{X}\, \underline{\beta}_2 = \underline{\beta}_3$$

Substitute in the LH side of eqn (4 - 3):

$$(\underline{\alpha}_3 - \underline{\beta}_3)^T \underline{\mathbf{x}} = (\underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 - \underline{\beta}_1 \mathbf{X} \underline{\beta}_2)^T k\,(\underline{\alpha}_1 - \underline{\beta}_1)\,\mathbf{X}\,(\underline{\alpha}_2 - \underline{\beta}_2)$$

$$= k\,(\underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 - \underline{\beta}_1 \mathbf{X} \underline{\beta}_2) \bullet (\underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 - \underline{\alpha}_1 \mathbf{X} \underline{\beta}_2 - \underline{\beta}_1 \mathbf{X} \underline{\alpha}_2 + \underline{\beta}_1 \mathbf{X} \underline{\beta}_2)$$

$$= k\,(\underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 \bullet \underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 - \underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 \bullet \underline{\alpha}_1 \mathbf{X} \underline{\beta}_2 - \underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 \bullet \underline{\beta}_1 \mathbf{X} \underline{\alpha}_2$$

$$+ \underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 \bullet \underline{\beta}_1 \mathbf{X} \underline{\beta}_2 - \underline{\beta}_1 \mathbf{X} \underline{\beta}_2 \bullet \underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2 + \underline{\beta}_1 \mathbf{X} \underline{\beta}_2 \bullet \underline{\alpha}_1 \mathbf{X} \underline{\beta}_2$$

$$+ \underline{\beta}_1 \mathbf{X} \underline{\beta}_2 \bullet \underline{\beta}_1 \mathbf{X} \underline{\alpha}_2 - \underline{\beta}_1 \mathbf{X} \underline{\beta}_2 \bullet \underline{\beta}_1 \mathbf{X} \underline{\beta}_2)$$

Now, remembering that $\underline{\mathbf{a}} \bullet \underline{\mathbf{b}}\,\mathbf{X}\,\underline{\mathbf{c}} = \underline{\mathbf{a}}\,\mathbf{X}\,\underline{\mathbf{b}} \bullet \underline{\mathbf{c}}$, this expression equals

$$k\,(1 - (\underline{\alpha}_1 \mathbf{X} \underline{\alpha}_2)\,\mathbf{X}\,\underline{\alpha}_1 \bullet \underline{\beta}_2 - \underline{\alpha}_1 \bullet \underline{\alpha}_2\,\mathbf{X}\,(\underline{\beta}_1 \mathbf{X} \underline{\alpha}_2)$$

$$+ 0 + \underline{\beta}_1 \bullet \underline{\beta}_2\,\mathbf{X}\,(\underline{\alpha}_1 \mathbf{X} \underline{\beta}_2) + (\underline{\beta}_1 \mathbf{X} \underline{\beta}_2)\,\mathbf{X}\,\underline{\beta}_1 \bullet \underline{\alpha}_2 - 1)$$

Now all the double cross product terms involve a repeated vector, so they are all zero. So we are left with:

$$(\underline{\alpha}_3 - \underline{\beta}_3)^T \underline{\mathbf{x}} = k(1 - 1) = 0$$

That is, eqn (4 - 3) is automatically satisfied if the first two parts of eqn (4) are, so eqn (4 - 3) is not an independent equation. *So a non-trivial* $\underline{\mathbf{x}}$ *exists for distinct systems* *(* $\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$*) and (* $\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$*)* . Call this vector $\underline{\mathbf{x}}_E$.

Now consider a rigid body in which we embed coordinates ($\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$) to orient it. We move the body in such a way that its point at the origin of coordinates is fixed. It now takes up the position where its reference axes lie along ($\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$). What is the significance of $\underline{\mathbf{x}}_E$? This point has the same coordinates in ($\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$) and ($\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$) i.e. it has not been altered by the body's move from ($\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$) to ($\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$) . The only way that this can be true for a rigid body is for $\underline{\mathbf{x}}_E$ to be an *axis* about which the body has rotated. So we have proved Euler's theorem:

The movement of a rigid body with one point fixed can be described by a (simple, single) rotation about some axis. (NB in the proof the fixed point is the origin).

This theorem is of profound importance in the theory of Plate Tectonics. This theory says that (large) regions of the Earth's solid outer shell (called the *lithospheric plates*) behave like rigid blocks. Their movement on the Earth's surface means that they have a fixed point - which is the centre of the Earth. Accordingly, the movement of one of these blocks relative to another, or relative to a fixed frame of reference, can be described as a rotation about some axis. This theorem is true for finite movements over large periods of geological time as well as for the current, instantaneous movement of the plates.

Index notation for coordinate transformations

Recall that we can write the three components of any vector $\underline{\mathbf{x}}$ in a new coordinate system by:

$$\underline{\mathbf{x}}' = A^T \underline{\mathbf{x}} \qquad (5)$$

where A = ($\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$) is the matrix whose columns are the vectors pointing along the new axes. Note that $\underline{\mathbf{x}}'$ is the *same* vector as $\underline{\mathbf{x}}$. What has changed, because the coordinate axes have changed, are the descriptions of its components.

We can write this equation in index notation as:

$$x'_i = \sum_{j=1}^{3} (A^T)_{ij} x_j$$

Write $a_{ij} = (A^T)_{ij}$. a_{ij} is the j th component of $\underline{\alpha}_i$.

Remember that according to our index set notation, a_{ij} means either an element of a matrix or the whole matrix.

Using the repeated index convention and dropping the Σ ,

$$x'_i = a_{ij} x_j \qquad (5)$$

Throughout these lectures we shall switch between the two notations of equations 5 as it suits us. The important thing is that *both equations represent the same set of three equations for the components of the LHS of eq 5.*

We can invert eqn (5) to get:

$$\underline{\mathbf{x}} = A\,\underline{\mathbf{x}}' \qquad (5a)$$

or in index notation:

$$x_i = A_{ij}\, x_j' = a_{ji}\, x_j' \qquad (5a)$$

which describes the coordinates of the vector x_i in the *old* (original) coordinate system in terms of those (x_j') in the *new* one.

Physical interpretation of transformation rules

Recall that we started our discussion of transformations by asking what the implications of the Fundamental Principle were for vectors. We have reached this conclusion: for a vector to represent a physical quantity, its components must obey equation (5) (or 5a) when we change coordinate systems. This conclusion will be the basis for our definition of Cartesian Tensors.

Concept of Continuum

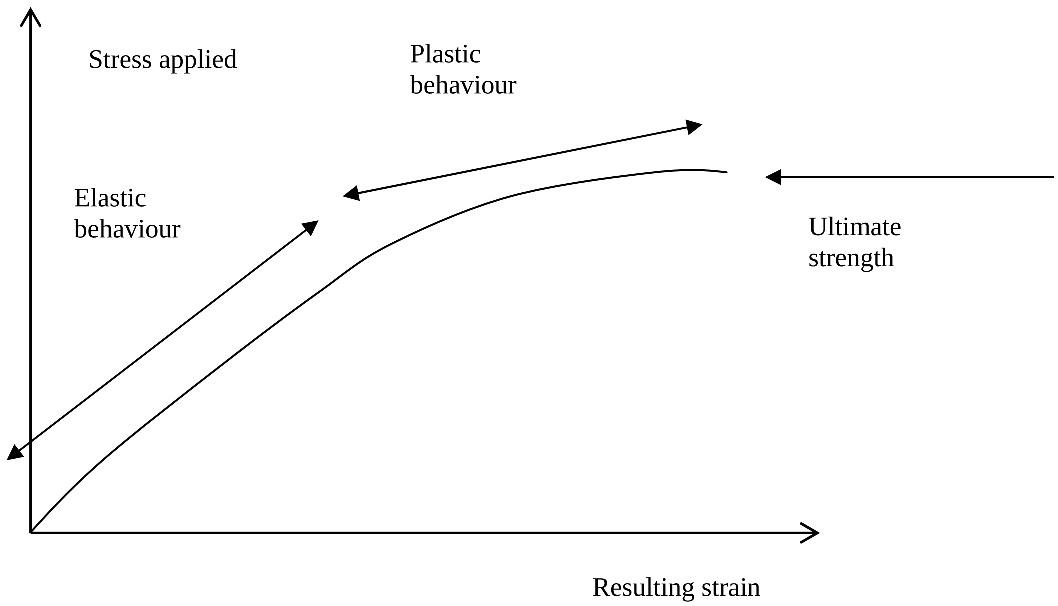
A *continuum* is a macroscopic model of a material: no atomic or quantum effects! The material is infinitely divisible and smoothly varying except, perhaps, for well defined points of discontinuity in the material properties or their gradients. The material may or may not be *homogeneous* – which means having the same properties at all points within the continuum, or within some sub-region of the continuum. It may or may not be *isotropic* – which means having the same properties in all directions at any point. These properties are independent. A plum pudding is isotropic but not homogeneous. A glass-fibre rod is homogeneous but not isotropic.

We will be dealing with *small deformations* of the continuum. Experiment shows that this assumption is usually necessary to invoke the assumption of *elastic behaviour* of the material. However, the methods can be extended (with some care) to finite deformations. And while we will be concerned with solids, much of the theory can be applied to fluids as well.

Rheology

It is worth briefly introducing the concept of a *rheological model* and some of the physical properties of continuous materials. To begin with, we are used to a number of terms that describe materials – *elastic, plastic, stiff, soft, fluid, brittle*, and so on. These terms describe qualitatively how a material reacts to stress.

A solid material's response to stress can be shown with a graph like this cartoon:



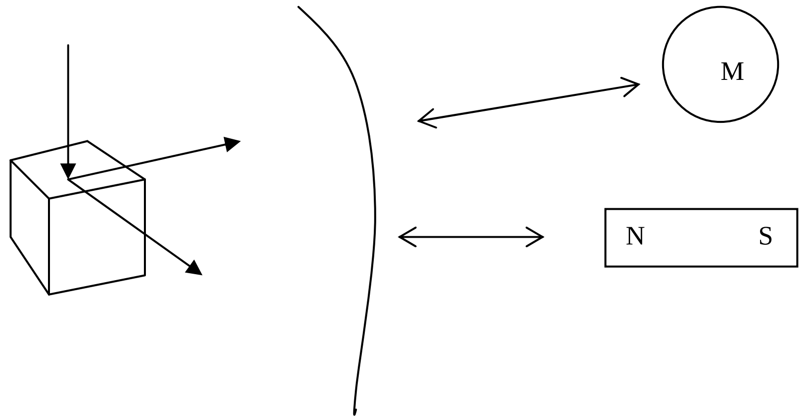
We can give these accurate, but still qualitative, meanings:

Elastic – the deformation is proportional to the magnitude of stress and the material's initial state is restored when the stress is removed.
Plastic – the deformation depends on the magnitude of the stress but also the time it is applied. If the stress is removed it will not return to its original state.
Stiff – the deformation per unit of stress is small
Soft – the deformation per unit of stress is large
Fluid – the *rate* of deformation depends the magnitude of the stress
(Ultimate) strength – the stress at which the (solid) material fractures
Brittle – a material with a small plastic region i.e the transition from elastic behaviour to failure is abrupt.

The formal quantitative relationship between stress and a material's deformation is called a *rheological model*. In this course we shall be working towards establishing one very common rheological model – that of an elastic solid.

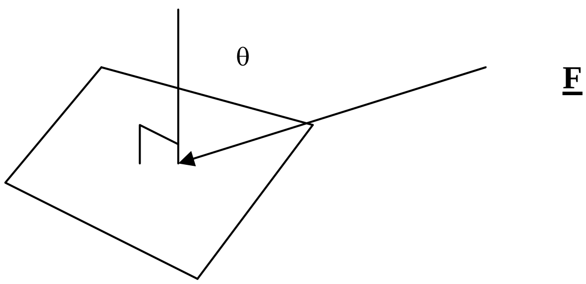
A physical quantity that is not a vector - stress

We wish to consider what forces act on a volume - a cuboid - of material within some body.

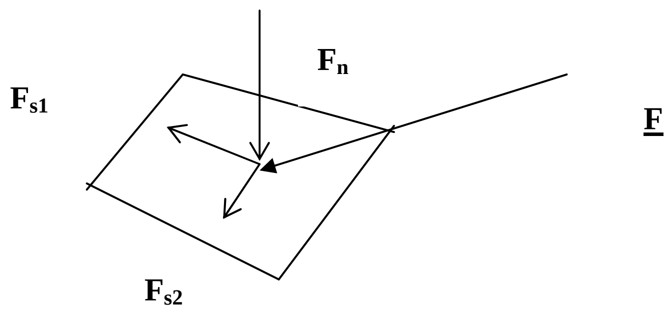


There are two classes of forces that can act on the cuboid. External forces, such as gravity or electromagnetic forces, which exert *bodily* on each particle of the cuboid, are called *body forces*, and the forces applied by the rest of the continuum on the cuboid, through its facets, which are *surface forces*. These surface forces will be described using the concept of *Stress*.

Force can be applied across the surface at any angle:



But we can resolve these perpendicular and parallel to the surface to give a normal forces and (two) tangential components:



F_n is the *normal* force and F_{s1} and F_{s2} are *shear* forces.

The force on a surface is likely to depend on the size of a surface. We need a measure of surface forces that is scale independent. Thinking of *pressure*, we can get this by dividing the forces by the area A of the surface. Then indeed the normal force may be replaced by normal force/area = pressure.

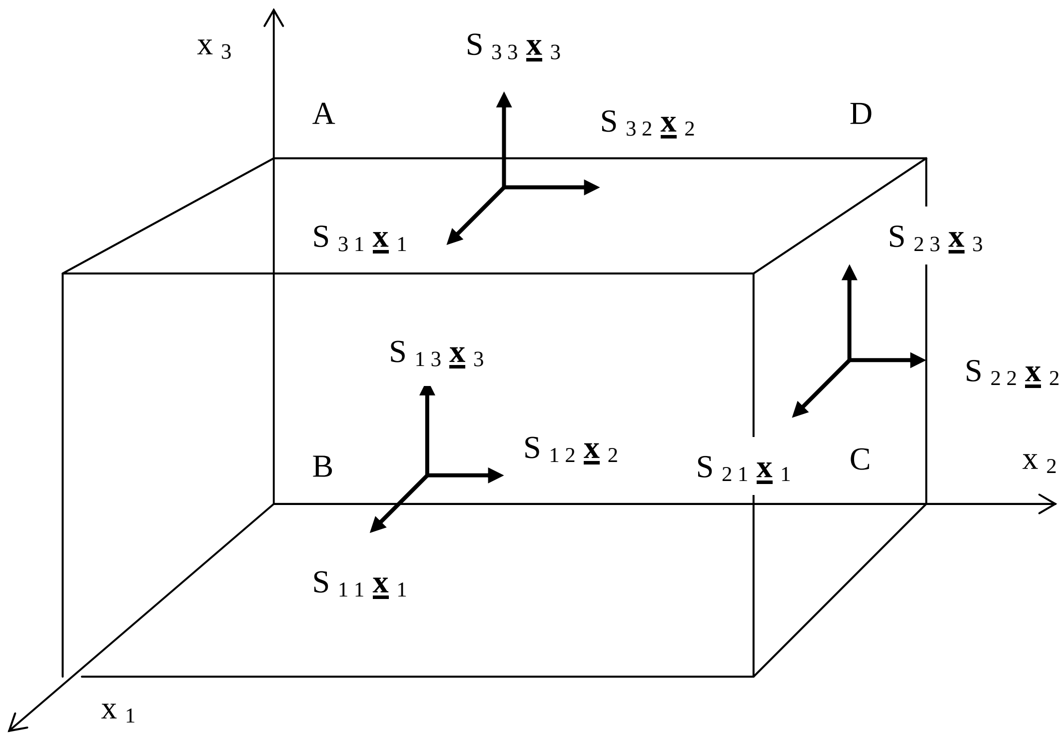
The two components of shear force are then transformed to *shear stresses*:

$S_1 = F_{s1} /A$
$S_2 = F_{s2} /A$

and we can do this for each of the six faces of the cuboid and so end up with (potentially) 6 x 3 = 18 components of stress.

Components of stress

We need a system for describing these. We will make things easy for ourselves by considering a cube (all faces with area A) that has its edges oriented along the coordinate axes:



1. The components of stress are depicted on the *outer* x_1x_3, x_2x_3, and x_1x_2 faces only. We will comment about the change in stress between parallel faces of the cube below.

2. The components of stress are indexed as follows; the *first* index is the index of the *normal to the face*; so on the x_2x_3 face we have three components of force per unit area: S_{11}, S_{12}, and S_{13}.

3. The second index gives the directions of the components on this face; so:

S_{11} is the component in the x_1 direction (on the x_2x_3 face)

S_{12} is the component in the x_2 direction (on the x_2x_3 face)

S_{13} is the component in the x_3 direction (on the x_2x_3 face)

4. *Sign convention*: We are considering forces *on* the cube. (The cube will of course exert equal and opposite forces back across each face on the rest of the continuum). Therefore:

If the outward normal to the face is in the +ve direction, then the stress components are $+ S_{ij}$;

If the normal is in the –ve direction, then the stress components are - S_{ij} .

Eg the *outward* normal to the x_2x_3 face is $\underline{\mathbf{x}}_1$; so we record the stress components on the cube across the x_2x_3 face as S_{1j} . The stress components on the parallel x_2x_3 face, which has outward normal - $\underline{\mathbf{x}}_1$, would be recorded as - S_{1j} .

[NB Imagine a similarly sized cube against, say, the x_2x_3 face. This would have an outward normal - $\underline{\mathbf{x}}_1$ for this face, so the stress components for this face would be - S_{1j} ; ie equal and opposite to the original cube.]

5. We could describe any one (e.g. i th) of the stress components, using index notation, by $S_{ij}\,\underline{\mathbf{x}}_j$. The corresponding force is stress x area = A $S_{ij}\,\underline{\mathbf{x}}_j$.

Now if we apply our index set rules to the expression $S_{ij}\,\underline{\mathbf{x}}_j$, we have to sum over the index j - i.e. $S_{ij}\,\underline{\mathbf{x}}_j$ means

$$\sum_{j=1}^{3} S_{ij}\,\underline{\mathbf{x}}_j = (S_{i1}\,\mathbf{x_1} + S_{i2}\,\mathbf{x_2} + S_{i3}\,\mathbf{x_3})$$

This is the *total* stress on the x_2x_3 face. In future, to avoid conflict with the summation convention we will suppress the use of the direction vectors $\underline{\mathbf{x}}_j$ and read S_{ij} as the stress component in direction j on face i .

6. On the hidden faces: Assume that the force field affecting the cube is sufficiently slowly varying in space (and smoothly varying i.e. differentiable) that we can write the components of stress at a point δx_k on a nearby, hidden face using a Taylor series expansion:

$$S_{ij}(\underline{\mathbf{x}}) = - \{ S_{ij}(\underline{\mathbf{x}} = \underline{\mathbf{x}}_0) + \sum_{k=1}^{3} \partial/\partial x_k\, S_{ij}(\underline{\mathbf{x}} = \underline{\mathbf{x}}_0)\, \delta x_k + \text{negligible higher order terms}\} \qquad (6)$$

or, using summation convention:

$$S_{ij}(\underline{\mathbf{x}}) = - \{S_{ij}(\underline{\mathbf{x}}_0) + \partial/\partial x_k\, S_{ij}(\underline{\mathbf{x}}_0)\, \delta x_k \} \qquad (6)$$

where we remember that because of the sign convention, and that the hidden faces' normals all point in the negative directions, these components of stress are *negative*.

Therefore, on the 'hidden' face parallel to the x_2x_3 face, $\underline{\boldsymbol{\delta}\mathbf{x}} = (-\delta, 0, 0)$, where δ is the length of a cube face, and

$$S_{1j}(\underline{\mathbf{x}}) = - S_{1j}(\underline{\mathbf{x}}_0) + \partial/\partial x_1\, S_{1j}(\underline{\mathbf{x}}_0)\, \delta$$

If we now let the cube shrink to zero, $\delta \to 0$, and in the limit

$$\lim_{\delta \to 0} S_{1j}(\underline{\mathbf{x}}) = - S_{1j}(\underline{\mathbf{x}}_0)$$

Similarly for the other faces. So the only difference between stress on parallel faces of an infinitesimal cube arises from the sign convention.

Putting all this together: we can define the stresses at any point in a continuum by a 3 x 3 array (matrix) of numbers:

$$S_{ij} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

Note that we cannot conveniently and naturally deal with the number of parameters we need to describe stress by using a lower dimensioned quantity i.e. a vector. We need to use a two-dimensional array. This array represents a physical quantity (or quantities) and so, like a vector, we wish it to be independent of the coordinate system in which we happen to be operating. What is the appropriate transformation rule for such a quantity?

Extension of the transformation rule to higher dimensions

The stress array can be thought of as a 'vector of vectors'. As we have seen, the i th row of S_{ij} is a stress vector describing the forces on face with normal $\underline{\mathbf{x}}_i$. If we wish to ensure that S_{ij}, treated as a vector of vectors, was independent of the coordinate system used to describe it, then each row vector would have to satisfy eqn (5).

$$s'_k = a_{kj} s_j \qquad (5^*)$$

where s_j is the i th *row* of S_{ij} (transposed). So we should write this as

$$s'_k = s_j a^T_{jk} \qquad (6)$$

We want, then, all the rows of S_{ij}, which are vector quantities in their own right, to obey eqn (5*). So:

$$S_{ik}' = S_{ij}\ a^T_{jk} \ \text{(i fixed)} \qquad (7)$$

Now write

$$S_{ij} = \underline{\mathbf{v}}_i$$

i.e. regard S_{ij} as a column vector whose components are row vectors. We want this to obey eqn (5) also. So:

$$\underline{\mathbf{v}}'_m = a_{mi}\underline{\mathbf{v}}_i = a_{mi} S_{ij} \qquad (8)$$

Each ‘component’ of $\underline{\mathbf{v}}_i$ is a row of S_{ij}. So we can rewrite eqn (8) as

$$S_{mj}' = a_{mi} S_{ij} \qquad \text{(j fixed)} \qquad (8a)$$

NB Remember the rules about indices in index equations. In eqn 8a, i is a dummy, by the summation convention. S_{mj}' are the components of S_{ij} in the new coordinate system.

This discussion suggests we should put equation 7 and 8 together:

$$\mathbf{S_{mk}' = a_{mi} S_{ij}\ a^T_{jk}} \qquad \mathbf{(9)}$$

or

$$\mathbf{S_{mk}' = a_{mi} a_{kj}\ S_{ij}} \qquad \mathbf{(9^*)}$$

This equation in fact gives the transformation rule for a two-dimensional quantity to be a physical quantity, consistent with eqn (5). Note:

(i) There is a double summation over i and j, both of which are dummies.
(ii) The order of the terms on the RHS is irrelevant in index notation. For although they represent 2-D entities (matrices), they also represent the (real) components of those entities, and real multiplication commutes.
(iii) S_{mk}' is the description of the components of S_{ij} in the new coordinate system (whose directions are the rows of a_{kj}).

Matrix interpretation

We can interpret equation (9) in terms of the matrices S (for stress S_{ij}) S' (in the new coordinates) and a = a_{kj}.

We need to rewrite (9) as:

$$S_{mk}' = a_{mi} S_{ij} a_{kj}$$

$$= a_{mi} S_{ij} a^T_{jk}$$

Now, put in the summation signs implied by the summation convention:

$$S_{mk}' = \sum_i \sum_j a_{mi} S_{ij} a^T_{jk}$$

And this is the rule for matrix multiplication; i.e.

$$S' = a S a^T \qquad (9a)$$

Which is the equivalent of eqn (9) in matrix notation.

Cartesian Tensors

We are now in a position to define *Cartesian Tensors*

1. We say that an indexed set $\{X_{ij}$, i = 1, 2, 3; j = 1, 2, 3$\}$ described in some Cartesian Coordinate system A_0 is a *Cartesian Tensor of rank 2* if *and only if* its components in some other coordinate system A_1 are given by:

$$X_{ij}' = a_{ik} a_{jm} X_{km}$$

Where *a* is the (orthogonal) matrix that describes the transformation from A_0 to A_1 .

2. (Thus inspired:) We say that an index set with N indices $\{X_{ijkm...n}$, all indices = 1, 2, 3$\}$ described in some Cartesian Coordinate system A_0 is a *Cartesian Tensor of rank N* (or just 'tensor' for short) *if and only if* its components in some other coordinate system A_1 are given by:

$$X_{ijkm...n}' = a_{ip} a_{jq} a_{kr} a_{ms} \ldots a_{nt} X_{pqrs...t}$$

(with N replicates of *a*) where *a* is the (orthogonal) matrix that describes the transformation from A_0 to A_1 .

3. Note that by the definition 2 a vector x_i is a tensor of rank 1 because we have already shown that:

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

4. A scalar is the same in all coordinate systems. This means that we can identify it as a tensor of rank 0.

The Kronecker tensor

A special tensor (of rank 2) is the Kronecker Tensor δ_{ij} defined by: $\delta_{ij} = 1$ for i = j, $\delta_{ij} = 0$, for $i \neq j$. That is, it has the matrix representation as I_{3x3} and is the *identity* for tensors of rank 2, as we can easily show:

Let X_{ij} be a tensor of rank 2. Then, remembering that

$$\delta_{ij} X_{jm} = \sum_{j=1}^{3} \delta_{ij} X_{jm}$$

$$\delta_{ij} X_{jm} = 1.\, X_{jm} \text{ (for } j = i) + 0.\, X_{jm} \text{ (for the two values of } j \neq i) = X_{im}$$

The Kronecker Tensor δ_{ij} is indeed a tensor i.e. it obeys the transformation rule for 2nd rank tensors. Proof – assignment.

Properties of tensors

1. A linear combination of two tensors of the same rank is also a tensor.

Proof: (Use rank 2 to demonstrate; extension to rank N is immediate). Let X_{ij}, Y_{ij} be two tensors of rank 2; therefore they transform according to:

$$X_{ij}' = a_{ik} a_{jm} X_{km}$$

$$Y_{ij}' = a_{ik} a_{jm} Y_{km}$$

and let α, β be two scalars. Then

$$a_{ik} a_{jm} (\alpha Y_{km} + \beta X_{km}) = \alpha a_{ik} a_{jm} Y_{km} + \beta a_{ik} a_{jm} X_{km}$$

$$= \alpha Y_{ij}' + \beta X_{ij}'$$

which means that $(\alpha Y_{k,m} + \beta X_{k,m})$ transforms as a tensor; QED.

2. The product of two tensors of rank N_1 and N_2 is a tensor of rank $N_1 + N_2$.

Eg: If X_{ij}, Y_{kmn} are two tensors of rank 2 and 3 respectively, they transform according to:

$$X_{ij}' = a_{ik} a_{jm} X_{km}$$

$$Y_{kmn}' = a_{kp} a_{mq} a_{nr} Y_{pqr}$$

Then $Z_{ijkmn} = X_{ij} Y_{kmn}$ transforms according to

$$a_{ik} a_{jm} a_{kp} a_{mq} a_{nr} Z_{ijkmn} = a_{ik} a_{jm} a_{kp} a_{mq} a_{nr} X_{ij} Y_{kmn}$$

$$= a_{ik} a_{jm} X_{ij} \, a_{kp} a_{mq} a_{nr} Y_{kmn}$$

$$= X_{ij}' \, Y_{kmn}'$$

$$= Z_{ijkmn}' \qquad \text{QED.}$$

3. If we sum over a pair of indices, it is called a *contraction.* A contraction is a tensor of rank 2 less than its parent.

eg: If Y_{kmn} is a tensor of rank 3 it transforms according to:

$$Y_{kmn}' = a_{kp} a_{mq} a_{nr} Y_{pqr}$$

Contract it over the second and third indices: $Y_k' = Y_{kmm}'$

Then: $Y_{kmm}' = a_{kp} a_{mq} a_{mr} Y_{pqr}$

$= a_{kp} a^T_{qm} a_{mr} Y_{pqr}$

Now remember that a is orthogonal, so $a^T a = I$; or in tensor notation:

$$a^T_{qm} a_{mr} = \delta_{qr}$$

So: $Y_{kmm}' = a_{kp} \delta_{qr} Y_{pqr}$

$= a_{kp} Y_{pqq}$

which is the transformation rule for a tensor of rank 1. QED.

4. *The derivative of a tensor is a tensor.* We assume that a tensor (rank 2) Y_{ij} is a differentiable function of position (x_1, x_2, x_3). The *gradient* of $Y_{ij} = \partial Y_{ij} / \partial x_k$ is a tensor of rank 3. To prove this we need to determine the transformation rule for a position vector $\underline{\mathbf{x}} = x_i$.

If *a* is the matrix describing the transformation from the original coordinate system to a new one, then we know that:

$$x_i' = a_{ik} x_k$$

or: $$x_j = a_{ij} x_i' \qquad (2b)$$

(NB $a^T a = I$ is that same as $a_{ij} a_{ik} = \delta_{jk}$)

So we can calculate:

$$\partial x_j / \partial x_i' = a_{ij} \qquad (10)$$

NB in eq 2 b, we sum over the index i. In eq 10 we pick out just the one multiplier of x_i' .

Y_{ij} is a tensor of rank 2 so:

$$Y_{ij}' = a_{in} a_{jm} Y_{nm}$$

Now calculate:

$\partial Y_{ij}' / \partial x_k'$ = (chain rule) $\partial Y_{ij}' / \partial x_r . \partial x_r / \partial x_k'$

(NB summation over r)

$= \partial (a_{in} a_{jm} Y_{nm})/ \partial x_r . a_{kr}$

$= a_{in} a_{jm} a_{kr} \partial Y_{nm} / \partial x_r$ (*a* is fixed)

ie $\partial Y_{ij}' / \partial x_k' = a_{in} a_{jm} a_{kr} \partial Y_{nm} / \partial x_r$

which is the required rule for $\partial Y_{nm} / \partial x_r$ to transform as a tensor of the 3rd rank. QED.

Notation: It is common to write the derivative of a tensor using a comma among the indices e.g. $\partial Y_{nm} / \partial x_r$ is written $Y_{nm,r}$. The gradient of a scalar *a* would be written $a_{,r}$.

Alternating tensor

Another special tensor is the alternating tensor, defined as:

ε_{ijk} = +1 if i, j, k are in cyclic order (e.g. 2,3,1)
= -1 if i, j, k are in anticyclic order (e.g. 1,3,2)
= 0 if two or more of i, j, k are equal (e.g. 2,2,1)

This enables vector algebra and vector calculus expressions to be written compactly in tensor notation. Eg for vectors **v** and **w**

$$\mathbf{u} = \mathbf{v} \times \mathbf{w} \text{ is written: } u_i = \varepsilon_{ijk} v_j w_k$$

For a 3x3 matrix A = a_{ij},

$$|A| = \varepsilon_{ijk} a_{1i} a_{2j} a_{3k}$$

We can use ε_{ijk} to write the vector derivative *curl:*

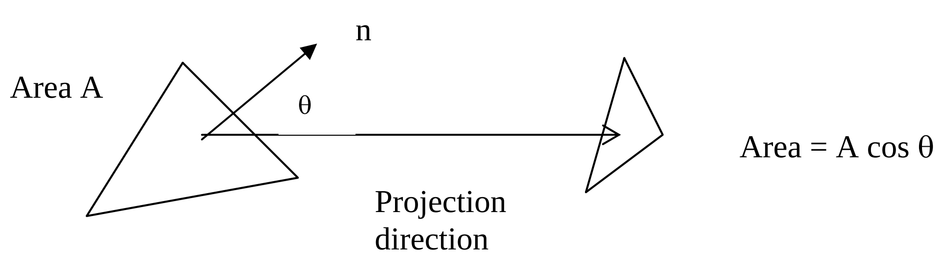
$$\text{curl } \mathbf{u} = \partial / \partial x_i \times \mathbf{u} \text{ is written: } (\text{curl } u)_i = \varepsilon_{ijk} (\partial / \partial x_j) u_k = \varepsilon_{ijk} \partial u_k / \partial x_j$$

Proving that stress is a tensor

We return now show formally that stress is a tensor, by showing that it obeys the transformation rule for tensors of the second rank.

Lemma

The projected area of a triangle = area of the triangle x cosine of the angle between the normal to the triangle and the projection direction; ie:



NB this is a very well known result, but one that is hardly ever proved!

Proof: let **a**, **b** be two sides of the original triangle. Then the area of the triangle, A, is given by:

$$A = \tfrac{1}{2} |\mathbf{a} \times \mathbf{b}|$$

And its (unit) normal

$$\mathbf{n} = \mathbf{a} \times \mathbf{b} / (|\mathbf{a} \times \mathbf{b}|)$$

Without loss of generality, project this in the x_1 direction. **a** can be written:

$$\mathbf{a} = a_1 \mathbf{x_1} + a_2 \mathbf{x_2} + a_3 \mathbf{x_3} \quad (= a_j \mathbf{x_j} \text{ - summation convention})$$

Then the projection of **a** , = $\mathbf{a_p}$ say, is given by:

$$\mathbf{a_p} = a_2 \mathbf{x_2} + a_3 \mathbf{x_3}$$

Similarly for **b**,

$$\mathbf{b}_{\mathbf{p}} = b_2 \mathbf{x}_2 + b_3 \mathbf{x}_3$$

The area of the projected triangle is thus:

$$½ | \mathbf{a}_{\mathbf{p}} \text{ x } \mathbf{b}_{\mathbf{p}} | = ½ | a_2 b_3 - a_3 b_2 |$$

and cos θ = projection direction • normal = (1, 0, 0) • **a** x **b** / (| **a** x **b** |)

$$= (a_2 b_3 - a_3 b_2) / (| \mathbf{a} \text{ x } \mathbf{b} |)$$

So the projected area = ½ cos θ | **a** x **b** |

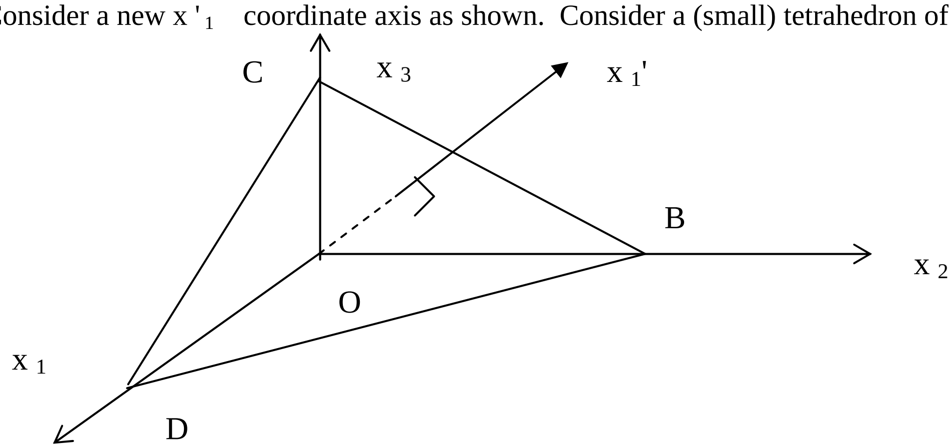
= cos θ x area of triangle QED.

Main result

Now we want to show that stress, as we have defined it, is a tensor (of rank 2); that is, that it satisfies a transformation rule:

$$S'_{ij} \text{ (new coords)} = a_{ip} a_{jq} S_{pq} \text{ (old coords)}$$

Consider a new x'_1 coordinate axis as shown. Consider a (small) tetrahedron of material OBCD:



Take arbitrary axes x'_2 and x'_3 at right angles to each other and x'_1. Let the area of BCD be δ. The cosines of the angles between x'_1 and x_1, x_2 and x_3 are a_{11}, a_{12}, and a_{13}, by the rule for the construction of matrix a_{ij}. So the areas of OBC, OCD and OBD are, by the Lemma,
δa_{11}, δa_{12}, and δa_{13} respectively.

So the stress *forces* acting *on* the tetrahedron are, using the summation convention (j):

$- S_{1j} \underline{\mathbf{x}}_{\mathbf{j}} \, \delta . a_{11}$ on OCB,

$- S_{2j} \underline{\mathbf{x}}_{\mathbf{j}} \, \delta . a_{12}$ on OCD,

$- S_{3j} \underline{\mathbf{x}}_{\mathbf{j}} \, \delta . a_{13}$ on ODB,

in the *old* coordinate system. NB negative signs because outward normals point in the opposite directions from the axes. Cf '*Components of stress*' paragraph 5.

And we can calculate the force on BCD in the *new* coordinate system to be:

$$S_{1i}' \underline{\mathbf{x}}_i' \delta$$

Now we want to calculate the forces on the OCB, OCD and ODB faces in the new system. We can do this by taking the scalar product of the force with $\underline{\mathbf{x}}_i'$. The scalar products of the original coordinate vectors $\underline{\mathbf{x}}_j$ with $\underline{\mathbf{x}}_i'$ are the direction cosines of $\underline{\mathbf{x}}_i'$, which, again, are the i th row of the transformation matrix a_{ij}.

So the components of force in the $\underline{\mathbf{x}}_i'$ direction are:

$$- S_{1j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{11}\, \delta \text{ on OCB,}$$

$$- S_{2j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{12}\, \delta \text{ on OCD,}$$

$$- S_{3j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{13}\, \delta \text{ on ODB,}$$

Now add up all the forces on the tetrahedron:

$$S_{1i}' \underline{\mathbf{x}}_i' \delta - S_{1j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{11}\, \delta - S_{2j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{12}\, \delta - S_{3j}\, a_{ij}\, \underline{\mathbf{x}}_i'\, a_{13}\, \delta$$

$$= \underline{\mathbf{x}}_i' (S_{1i}' - S_{kj}\, a_{ij}\, a_{1k})\, \delta \quad \text{(summation convention for k)}$$

That is, there is a stress force component in the direction $\underline{\mathbf{x}}_i'$ of

$$(S_{1i}' - S_{kj}\, a_{ij}\, a_{1k})\, \delta$$

Call this $T_i\ \delta$.

If there are body forces G_i per unit mass, and if the density and volume of the tetrahedron are ρ and V, then Newton's Law of motion is:

$$\rho V\, \partial^2 \underline{\mathbf{x}}_i' / \partial t^2 = \rho V\, G_i + T_i\ \delta$$

Divide by δ and let the size of the tetrahedron ~ $V / \delta \to 0$ ie:

$$\rho V/\delta\ \partial^2 \underline{\mathbf{x}}_i' / \partial t^2\ (\to 0) = \rho V/\delta\ G_i\ (\to 0) + T_i$$

Therefore in the limit as the tetrahedron → 0, $T_i = 0$; ie:

$$(S_{1i}' - S_{kj}\, a_{ij}\, a_{1k}) = 0 \quad \text{or}$$

$$S_{1i}' = a_{1k}\, a_{ij}\, S_{kj}$$

Now repeat this for axes 2, 3 and we have:

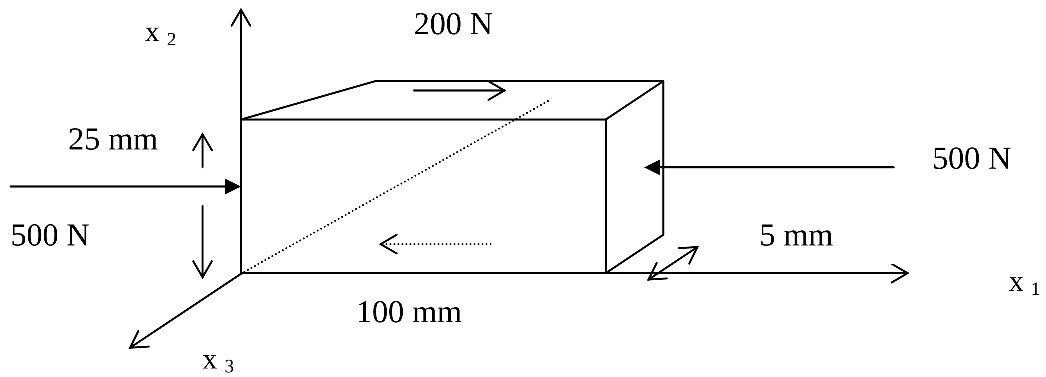
$$S_{mi}' = a_{mk}\, a_{ij}\, S_{kj}$$

Which is the required transformation rule; i.e. S_{kj} is a tensor.

NB: We have showed that S_{kj} being a tensor follows from the requirement that the continuum obeys Newton's Law of motion.

Example – Calculation of the Stress Tensor

A block of metal is clamped in its long direction (x_2x_3 faces) with a force of 500 N and then shear forces of 200 N are applied to the x_1x_3 faces.



(i) What shear forces must the clamps apply to the (x_2x_3) faces to maintain equilibrium?

(ii) Assuming that stresses are uniform throughout the block, what is the Stress Tensor?

Answer

(i) There must be an anti-clockwise moment on the x_2x_3 faces to offset the clockwise moment of the applied shears. If the force producing this moment is F, taking moments about the centre of the block gives:

$$- (2 \times 200 \times 25/2) \text{ N mm} + (2 \times F \times 100/2) \text{ N mm} = 0 \text{ for equilibrium}$$

$$\therefore \; F = 50 \text{ N}$$

(ii) The components of the stress tensor S_{ij} are:

x_1 (x_2x_3) face: $S_{11} = -500 / (25 \times 5) \text{ N / mm}^2 = -4 \text{ N / mm}^2$

[NB: it is -ve because the 500 N acts in the opposite direction to the outward normal]

$$S_{12} = +F / (25 \times 5) \text{ N / mm}^2 = 50 / 125 = 0.4 \text{ N / mm}^2$$

[NB: it is +ve because e.g. on the LH face, F acts in the - ve **x_2** direction and the outward normal is – **x_1**]

$$S_{13} = 0$$

x_2 face:

$$S_{21} = -(-200) / (100 \times 5) \text{ N / mm}^2 = 0.4 \text{ N / mm}^2$$

$$S_{22} = 0$$

$$S_{23} = 0$$

x_3 face:

$$S_{31} = 0$$

$$S_{32} = 0$$

$$S_{33} = 0$$

$$\text{so } S = \begin{bmatrix} -4 & 0.4 & 0 \\ 0.4 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad N / mm^2$$

Note that S is *symmetric* – this is a consequence of balancing the moments of the shears for equilibrium, and is a general result, as we shall see.

Stress force across an arbitrary (plane) surface

We can use the stress tensor to calculate the components of force across an arbitrary plane in the continuum. We make use of the result from the analysis of the tetrahedron (which showed that stress was a tensor).

Without loss of generality, take the angled face of the tetrahedron (with normal x_1') to be the arbitrary plane we seek of calculate the forces across.

If (as before) the area of the plane is δ, we had the components of the stress force across the plane in the new (ie ') coordinate system being:

$$S_{1i}'\,\delta = a_{ij}\, a_{1k}\, S_{kj}\, \delta$$

or

$$S_{1i}' = a_{ij}\, a_{1k}\, S_{kj} \qquad \text{per unit area.}$$

Write:

$$F_i'^{\,T} = S_{1i}' = a_{ij}\, a_{1k}\, S_{kj} = a_{1k}\, S_{kj}\, (a^T)_{ji}$$

And note that $F_i'^{\,T}$ is a row vector; so:

$$F_i' = (a_{1k}\, S_{kj}\, (a^T)_{ji}\ )^T$$

$$= a_{ij}\, S_{jk}\, (a^T)_{k1}$$

Now F_i' is a (column) vector, so it transforms according to the rule:

$$F_i' = a_{ij}\, F_j$$

Where F_j are its components in the old (x_1 , x_2 , x_3) coordinate system.

So:

$$F_j = S_{jk}\, (a^T)_{k1}$$

But the components of a_{1k} are, by construction, the cosines of the angle between (new x_1 , old x_k), which is the *normal* to the plane (in old coordinates).

So:

$$a_{1k} = n_k^T$$

And so:

Or in matrix notation:

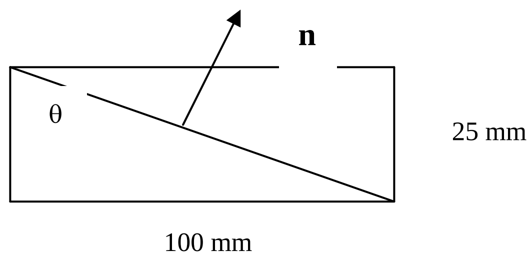
$$F_j = S_{jk}\, n_k$$

$$\mathbf{F} = \mathbf{S}\, \underline{\mathbf{n}}$$

is the equation that gives the stress force per unit area across a plane surface with normal **n** = n_k

Example:

In the block of metal examined earlier, what is the stress force per unit area across the plane in the block that runs from corner to corner in the x_1x_2 face?



$n_1 = \cos(\tan^{-1} 100/25) = 0.2425$
$n_2 = \sin(\tan^{-1} 100/25) = 0.9701$
$n_3 = 0$

So:

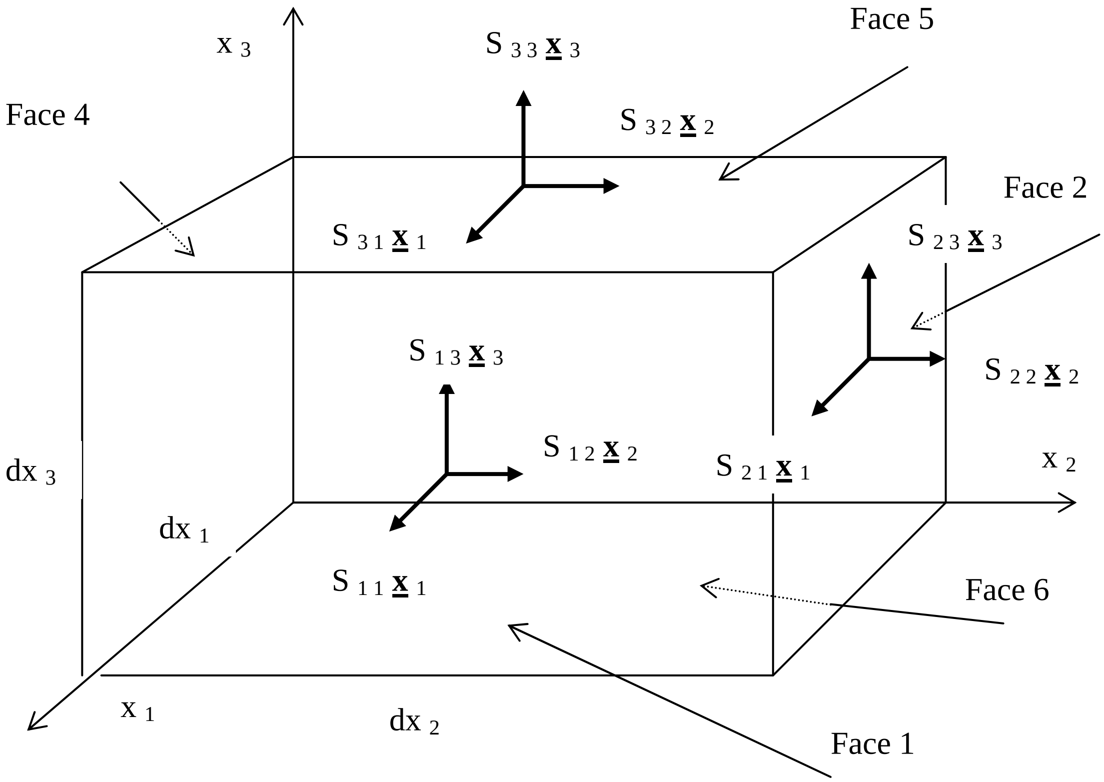
$$\mathbf{F} = \begin{bmatrix} -4 & 0.4 & 0 \\ 0.4 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.2425 \\ 0.9701 \\ 0.0 \end{bmatrix} = \begin{bmatrix} -0.5820 \\ 0.0970 \\ 0.0 \end{bmatrix} \quad \text{MN / m}^2$$

Thus the normal stress force per unit area = $\mathbf{F} \bullet \mathbf{n}$ (or $F_i\, n_i$) = - 0.0470 MN / m^2

And the shear stress = (Total ^{2} - Normal ^{2}) ^{½} = 0.588 MN / m^2

Symmetry of the Stress Tensor

The Stress Tensor is (always) symmetric. As we noted before, this is a consequence of the need for the shear force moments to cancel when a particle within a material is in equilibrium. Here is an outline of the proof.



See fig above. NB Face 3 is opposite Face 4.

If S_{1j} are the stress components on the x_2x_3 face of the cube $d\,x_1\,d\,x_2\,d\,x_3$ (Face 1), then on the parallel face, Face 2, to the first order in $d\,x_1$ the stress will be:

$$- (S_{1j} + \partial S_{1j} / \partial x_1 \; d\,x_1)$$

Similarly, the stress on the faces Face 3 and Face 4 are

$$S_{2j} \quad \text{and} \quad - (S_{2j} + \partial S_{2j} / \partial x_2 \; d\,x_2)$$

And the stress on the faces Face 5 and Face 6 are

$$S_{3j} \quad \text{and} \quad - (S_{3j} + \partial S_{3j} / \partial x_3 \; d\,x_3)$$

Now take moments (anticlockwise +ve) of all the stress *forces* from these stresses (ie multiply by the areas) about an axis parallel to x_3 through the centre of the cuboid.

For this axis, the forces on Faces 5 & 6 have no moment, as they act perpendicular or parallel to the axis.

Similarly, the moments of the forces of S_{22} and S_{23} (on Face 3), and the corresponding forces on Face 4, and S_{11} and S_{13} on Face 1, and the corresponding forces on Face 2, have no moment as they act either perpendicular or parallel to x_3.

So the contribution to the moments about the axis come from:

S_{21} on Face 3: anticlockwise moment

$$= - S_{21} \text{ (area) x (1/2 distance to axis)}$$

$$= - S_{21}\ d x_1\, d x_3\, .\, d x_2 /2$$

equivalent on Face 4: anticlockwise moment

$$= + (- (S_{21} + \partial S_{21} /\partial x_2\ d x_2))\ d x_1\, d x_3\, .\, d x_2 /2$$

S_{12} on Face 1: anticlockwise moment

$$= + S_{12}\ d x_2\, d x_3\, .\, d x_1 /2$$

equivalent on Face 2: anticlockwise moment

$$= - (- (S_{12} + \partial S_{12} /\partial x_1\ d x_1))\ d x_2\, d x_3\, .\, d x_1 /2$$

Adding, the total moment is:

$$- 2 S_{21}\, d x_1\, d x_3\, .\, d x_2 /2 - \partial S_{21} /\partial x_2\ d x_1\, d x_3\, .\, d x_2{}^2/2$$
$$+ 2 S_{12}\, d x_2\, d x_3\, .\, d x_1 /2 + \partial S_{12} /\partial x_1\ d x_2\, d x_3\, .\, d x_1{}^2/2$$

This moment provides a torque that causes an angular acceleration α
to the mass ($\rho\, d x_1\, d x_2\, d x_3$) of the cuboid, ie

$$- 2 S_{21}\, d x_1\, d x_3\, .\, d x_2 /2 - \partial S_{21} /\partial x_2\ d x_1\, d x_3\, .\, d x_2{}^2/2$$
$$+ 2 S_{12}\, d x_2\, d x_3\, .\, d x_1 /2 + \partial S_{12} /\partial x_1\ d x_2\, d x_3\, .\, d x_1{}^2/2 = \alpha\, \rho\, d x_1\, d x_2\, d x_3$$

Divide through by the volume $d x_1\, d x_2\, d x_3$ of the cuboid, and let $d x_j \to 0$. Then

$$- 2 S_{21} /2 - \partial S_{21} /\partial x_2\, .\, d x_2 /2 + 2 S_{12} /2 + \partial S_{12} /\partial x_1\, .\, d x_1 /2 = \alpha\, \rho$$

and in the limit as $d x_j \to 0$,

$$- S_{21} + S_{12} = \alpha\, \rho$$

So if the cuboid is in equilibrium, ie $\alpha = 0$, then:

$$- S_{21} + S_{12} = 0 \text{ or } S_{21} = S_{12}$$

Now apply this method to the moments about the other axes x_1 and x_2 , and we have, in turn, that

$$S_{23} = S_{32}$$
$$S_{13} = S_{31}$$

ie S is symmetric. QED.