Chapter 3

Cartesian Tensors

3.1 Suffix Notation and the Summation Convention

We will consider vectors in 3D, though the notation we shall introduce applies (mostly) just as well to n dimensions. For a general vector

$$\mathbf{x} = (x_1, x_2, x_3)$$

we shall refer to x_i, the i^{th} component of $\mathbf{x}$. The index i may take any of the values 1, 2 or 3, and we refer to "the vector x_i" to mean "the vector whose components are (x_1, x_2, x_3)". However, we cannot write $\mathbf{x} = x_i$, since the LHS is a vector and the RHS a scalar. Instead, we can write $[\mathbf{x}]_i = x_i$, and similarly $[\mathbf{x} + \mathbf{y}]_i = x_i + y_i$.

Note that the expression $y_i = x_i$ implies that $\mathbf{y} = \mathbf{x}$; the statement in suffix notation is implicitly true for all three possible values of i (one at a time!).

Einstein introduced a convention whereby if a particular suffix (e.g., i) appears twice in a single term of an expression then it is implicitly summed. For example, in traditional notation

$$\mathbf{x} \,.\, \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 = \sum_{i=1}^{3} x_i y_i;$$

using summation convention we simply write

$$\boxed{\mathbf{x} \,.\, \mathbf{y} = x_i y_i.}$$

All we are doing is not bothering to write down the $\sum$!

The Rules of Summation Convention

Summation convention does not allow any one suffix to appear more than *twice* within a single term; so $x_i y_i z_i$ is meaningless. We have to take care to avoid this: for example, consider the vector relation

$$\mathbf{y} = (\mathbf{a} \mathbin{.} \mathbf{b})\mathbf{x}.$$

We have $\mathbf{a} \mathbin{.} \mathbf{b} = a_i b_i$, but we *cannot* write $y_i = a_i b_i x_i$ as this would be ambiguous. How can we correct this? Note that

$$\mathbf{a} \mathbin{.} \mathbf{b} = a_i b_i = a_j b_j$$

– the suffix we use for the summation is immaterial. (Compare with the use of dummy variables in integrations: $\int_0^\infty e^{-x}\,\mathrm{d}x = \int_0^\infty e^{-y}\,\mathrm{d}y$.) So we *can* write

$$y_i = a_j b_j x_i.$$

In any given term, then, there are two possible types of suffix: one that appears precisely once, e.g., i in $a_j b_j x_i$, which is known as a *free suffix*; and one that appears precisely twice, e.g., j in $a_j b_j x_i$, which is known as a *dummy suffix*. It is an important precept of summation convention that the free suffixes must match precisely in every term (though dummy suffixes can be anything you like so long as they do not clash with the free suffixes). So in the equation

$$a_j b_j z_k = x_k + a_i a_i y_k b_j b_j$$

every term has a free suffix k, and all other suffixes are dummy ones. In vector notation, this equation reads

$$(\mathbf{a} \mathbin{.} \mathbf{b})\mathbf{z} = \mathbf{x} + |\mathbf{a}|^2 |\mathbf{b}|^2 \mathbf{y}.$$

(Note that the order of variables in the final term of this equation in suffix notation is unimportant: we could equally well have written $b_j y_k a_i b_j a_i$.)

There need not be any free suffixes at all, as in the equation $a_i z_i = (x_i + y_i) a_i$ (which reads $\mathbf{a} \mathbin{.} \mathbf{z} = (\mathbf{x} + \mathbf{y}) \mathbin{.} \mathbf{a}$ in vector notation).

Suffix notation can also be used with matrices. For a matrix A, we write a_{ij} to denote the entry in the i^{th} row and j^{th} column of A (for each $i = 1, 2, 3$ and $j = 1, 2, 3$). We write either $A = (a_{ij})$ or $[A]_{ij} = a_{ij}$ – these equations are equivalent – to indicate this. (Sometimes the upper-case letter is used instead, in which case the matrix A would have entries A_{ij}.)

Examples of Summation Convention

(i) $2\mathbf{x} + \mathbf{y} = \mathbf{z} \Leftrightarrow 2x_i + y_i = z_i$. Note that the RHS of this suffix notation equation does *not* mean $z_1 + z_2 + z_3$ – no repeated suffix, no sum!

(ii) $(\mathbf{a} \mathbin{.} \mathbf{b})(\mathbf{x} \mathbin{.} \mathbf{y}) = 0 \Leftrightarrow a_i b_i x_j y_j = 0$.

(iii) In summation convention, $\mathbf{y} = A\mathbf{x}$ is written

$$\boxed{y_i = [A\mathbf{x}]_i = a_{ij} x_j}$$

(check that this is correct by writing it out long-hand for each possible value of the free suffix i).

(iv) The matrix multiplication $C = AB$ (where A and B are 3×3 matrices) is written

$$\boxed{c_{ij} = [AB]_{ij} = a_{ik} b_{kj}.}$$

(v) The *trace* of a matrix C may be written as $\operatorname{Tr} C = c_{ii}$, i.e., $c_{11} + c_{22} + c_{33}$. Hence

$$\operatorname{Tr}(AB) = a_{ik} b_{ki}.$$

Replacing two free suffixes (e.g. i, j in c_{ij}) by a single dummy suffix (c_{ii}) is known as *contraction*.

Not all expressions written in suffix notation can be recast in vector or matrix notation. For example, $a_{ijk} = x_i y_j z_k$ is a valid equation in suffix notation (each term has three free suffixes, i, j and k), but there is no vector equivalent.

3.2 The Kronecker Delta and the Alternating Tensor

The Kronecker delta is defined by

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

and the alternating tensor is defined by

$$\boxed{\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is a cyclic permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an anti-cyclic permutation of } (1,2,3) \\ 0 & \text{if any of } i,\ j,\ k \text{ are equal} \end{cases}}$$

(i.e., $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1; \epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$; all others are zero). Note that $\delta_{ij} = \delta_{ji}$ and that $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik}$ etc.

If I is the identity matrix $\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ then $[I]_{ij} = \delta_{ij}$. We see that

$$x_i = \delta_{ij} x_j$$

because (i) this is equivalent to $\mathbf{x} = I\mathbf{x}$; or (ii) we can check for each value of i (e.g., when $i = 1$, RHS $= \delta_{1j}x_j = \delta_{11}x_1 + \delta_{12}x_2 + \delta_{13}x_3 = x_1 =$ LHS). The Kronecker delta just "selects" entries: e.g., $\delta_{ik}a_{jk}$ is equal to a_{ji}.

What is δ_{ii}? It is *not* 1.

The alternating tensor can be used to write down the vector equation $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ in suffix notation:

$$\boxed{z_i = [\mathbf{x} \times \mathbf{y}]_i = \epsilon_{ijk}x_j y_k.}$$

(Check this: e.g., $z_1 = \epsilon_{123}x_2y_3 + \epsilon_{132}x_3y_2 = x_2y_3 - x_3y_2$, as required.) There is one very important property of ϵ_{ijk}:

$$\boxed{\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.}$$

This makes many vector identities easy to prove.

(The property may be proved by first proving the generalisation

$$\epsilon_{ijk}\epsilon_{lmn} = \det \begin{pmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{pmatrix}.$$

Both sides clearly vanish if any of i, j, k are equal; or if any of l, m, n are. Now take $i = l = 1$, $j = m = 2$, $k = n = 3$: both sides are clearly 1. Finally consider the effect of swapping say i and j. Once we have proved this generalisation, contract k and n and simplify, noting that for example $\delta_{jk}\delta_{km} = \delta_{jm}$.)

Example: prove that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

$$\begin{aligned} [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk}a_j[\mathbf{b} \times \mathbf{c}]_k \\ &= \epsilon_{ijk}a_j\epsilon_{klm}b_lc_m \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})a_jb_lc_m \\ &= a_jb_ic_j - a_jb_jc_i \\ &= (\mathbf{a} \cdot \mathbf{c})b_i - (\mathbf{a} \cdot \mathbf{b})c_i \\ &= [(\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}]_i\,, \end{aligned}$$

as required.

ϵ_{ijk} can also be used to calculate determinants. The determinant of a 3×3 matrix $A = (a_{ij})$ is given by $\epsilon_{ijk}a_{1i}a_{2j}a_{3k}$ (check this by just expanding the product and sum in full). This can be written in several other ways; for example,

$$\det A = \epsilon_{ijk} a_{1i} a_{2j} a_{3k} = \epsilon_{jik} a_{1j} a_{2i} a_{3k}$$

[swapping i and j]

$$= -\epsilon_{ijk} a_{2i} a_{1j} a_{3k}.$$

This proves that swapping two rows of a matrix changes the sign of the determinant.

3.3 What is a Vector?

A vector is more than just 3 real numbers. It is also a physical entity: if we know its 3 components with respect to one set of Cartesian axes then we know its components with respect to any other set of Cartesian axes. (The *vector* stays the same even if its components do not.)

For example, suppose that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a right-handed orthogonal set of unit vectors, and that a vector $\mathbf{v}$ has components v_i relative to axes along those vectors. That is to say,

$$\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3 = v_j \mathbf{e}_j.$$

What are the components of $\mathbf{v}$ with respect to axes which have been rotated to align with a different set of unit vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$? Let

$$\mathbf{v} = v'_1 \mathbf{e}'_1 + v'_2 \mathbf{e}'_2 + v'_3 \mathbf{e}'_3 = v'_j \mathbf{e}'_j.$$

Now $\mathbf{e}'_i \cdot \mathbf{e}'_j = \delta_{ij}$, so

$$\mathbf{v} \cdot \mathbf{e}'_i = v'_j \mathbf{e}'_j \cdot \mathbf{e}'_i = v'_j \delta_{ij} = v'_i$$

but also

$$\mathbf{v} \cdot \mathbf{e}'_i = v_j \mathbf{e}_j \cdot \mathbf{e}'_i = v_j l_{ij}$$

where we define the matrix $L = (l_{ij})$ by

$$\boxed{l_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j.}$$

Then

$$\boxed{v'_i = l_{ij} v_j}$$

(or, in matrix notation, $\mathbf{v}' = L\mathbf{v}$ where $\mathbf{v}'$ is the column vector with components v'_i). L is called the *rotation matrix*.

This *looks* like, but is *not quite the same as*, rotating the vector $\mathbf{v}$ round to a different vector $\mathbf{v}'$ using a transformation matrix L. In the present case, $\mathbf{v}$ and $\mathbf{v}'$ are the *same* vector, just measured with respect to different axes. The transformation matrix corresponding to the rotation $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \mapsto \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is not L (in fact it is L^{-1}).

Now consider the reverse of this argument. Exactly the same discussion would lead to

$$v_i = \hat{l}_{ij} v'_j$$

where

$$\hat{l}_{ij} = \mathbf{e}_i \cdot \mathbf{e}'_j$$

(we swap primed and unprimed quantities throughout the argument). We note that $\hat{l}_{ij} = l_{ji}$ from their definitions; hence

$$\hat{L} = L^T$$

and so

$$\mathbf{v} = \hat{L}\mathbf{v}' = L^T \mathbf{v}'.$$

We can deduce that

$$\mathbf{v} = L^T L \mathbf{v},$$

and furthermore, this is true for *all* vectors $\mathbf{v}$. We conclude that

$$L^T L = I,$$

i.e.,

$$\boxed{L^T = L^{-1}.}$$

(Hence $LL^T = I$ also.) L is therefore an orthogonal matrix. In suffix notation, the equation $L^T L = I$ reads

$$l_{ki} l_{kj} = \delta_{ij},$$

and $LL^T = I$ reads

$$l_{ik} l_{jk} = \delta_{ij};$$

both of these identities will be useful.

Another way of seeing that $LL^T = I$ (or, equivalently, $L^T L = I$) is to consider the components of L. Since $\mathbf{e}'_i \cdot \mathbf{e}_j$ is just the j^{th} component of $\mathbf{e}'_i$ measured with respect to the first frame, we see that the i^{th} row of L just consists of the components of $\mathbf{e}'_i$ measured with respect to the first frame:

$$L = \begin{pmatrix} \mathbf{e}'_1 \cdot \mathbf{e}_1 & \mathbf{e}'_1 \cdot \mathbf{e}_2 & \mathbf{e}'_1 \cdot \mathbf{e}_3 \\ \mathbf{e}'_2 \cdot \mathbf{e}_1 & \mathbf{e}'_2 \cdot \mathbf{e}_2 & \mathbf{e}'_2 \cdot \mathbf{e}_3 \\ \mathbf{e}'_3 \cdot \mathbf{e}_1 & \mathbf{e}'_3 \cdot \mathbf{e}_2 & \mathbf{e}'_3 \cdot \mathbf{e}_3 \end{pmatrix}$$

$$= \begin{pmatrix} {\mathbf{e}'_1}^T \\ \hline {\mathbf{e}'_2}^T \\ \hline {\mathbf{e}'_3}^T \end{pmatrix} \qquad \text{[measured with respect to frame 1].}$$

Alternatively, the i^{th} column consists of the components of $\mathbf{e}_i$ with respect to the second frame.

To calculate the top left component of LL^T, we find the dot product of the first row of L with the first column of L^T. Both are simply $\mathbf{e}'_1$ measured with respect to the first frame, so we obtain $\mathbf{e}'_1 \cdot \mathbf{e}'_1$, which is 1. Similarly, the top right component of LL^T is $\mathbf{e}'_1 \cdot \mathbf{e}'_3$, which is zero. So, considering all possible combinations of rows and columns, we see that $LL^T = I$ as required.

3.4 Tensors

Tensors are a generalisation of vectors. We think informally of a tensor as something which, like a vector, can be measured component-wise in any Cartesian frame; and which also has a physical significance independent of the frame, like a vector.

Physical Motivation

Recall the conductivity law, $\mathbf{J} = \sigma\mathbf{E}$, where $\mathbf{E}$ is the applied electric field and $\mathbf{J}$ is the resulting electric current. This is suitable for simple isotropic media, where the conductivity is the same in all directions. But a matrix formulation may be more suitable in anisotropic media; for example,

$$\mathbf{J} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{E}$$

might represent a medium in which the conductivity is high in the x-direction but in which no current at all can flow in the z-direction. (For instance, a crystalline lattice structure where vertical layers are electrically insulated.)

More generally, in suffix notation we have

$$J_i = \sigma_{ij} E_j$$

where σ is the *conductivity tensor*.

What happens if we measure $\mathbf{J}$ and $\mathbf{E}$ with respect to a different set of axes? We would expect the matrix σ to change too: let its new components be σ'_{ij}. Then

$$J'_i = \sigma'_{ij} E'_j.$$

But $\mathbf{J}$ and $\mathbf{E}$ are vectors, so

$$J'_i = l_{ij} J_j$$

and

$$E_i = l_{ji} E'_j$$

from the results regarding the transformation of vectors in §3.3. Hence

$$\begin{aligned}
\sigma'_{ij} E'_j &= J'_i \\
&= l_{ip} J_p \\
&= l_{ip} \sigma_{pq} E_q \\
&= l_{ip} \sigma_{pq} l_{jq} E'_j \\
\implies \qquad (\sigma'_{ij} - l_{ip} l_{jq} \sigma_{pq}) E'_j &= 0.
\end{aligned}$$

This is true *for all* vectors $\mathbf{E}'$, and hence the bracket must be identically zero; hence $\sigma'_{ij} = l_{ip} l_{jq} \sigma_{pq}$. This tells us how σ transforms.

Compare this argument with the corresponding argument for the case $A\mathbf{x} = \mathbf{0}$ where A is a matrix; if it is true for all $\mathbf{x}$ then A must be zero, though this is not the case if it is only true for some $\mathbf{x}$'s.

σ is a *second rank* tensor, because it has two suffixes (σ_{ij}).

Definition: In general, a tensor of rank n is a mathematical object with n suffixes, $T_{ijk\ldots}$, which obeys the *transformation law*

$$\boxed{T'_{ijk\ldots} = l_{ip} l_{jq} l_{kr} \ldots T_{pqr\ldots}}$$

where L is the rotation matrix between frames.

Note: for second rank tensors such as σ, the transformation law

$$T'_{ij} = l_{ip} l_{jq} T_{pq}$$

can be rewritten in matrix notation as $T' = LTL^T$ – check this yourself!

Examples of Tensors

(i) Any vector $\mathbf{v}$ (e.g., velocity) is a tensor of rank 1, because $v'_i = l_{ip} v_p$.

(ii) Temperature T is a tensor of rank 0 – known as a *scalar* – because it is the same in all frames ($T' = T$).

(iii) The inertia tensor. Consider a mass m which is part of a rigid body, at a location $\mathbf{x}$ within the body. If the body is rotating with angular velocity $\boldsymbol{\omega}$ then the mass's velocity is $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{x}$, and its angular momentum is therefore

$$m\mathbf{x} \times \mathbf{v} = m\mathbf{x} \times (\boldsymbol{\omega} \times \mathbf{x}) = m\big(|\mathbf{x}|^2 \boldsymbol{\omega} - (\boldsymbol{\omega} \,.\, \mathbf{x})\mathbf{x}\big).$$

Changing from a single mass m to a continuous mass distribution with density $\rho(\mathbf{x})$, so that an infinitesimal mass element is $\rho(\mathbf{x})\,\mathrm{d}V$, we see that the total angular momentum of a rigid body V is given by

$$\mathbf{h} = \iiint\limits_V \rho(\mathbf{x})\big(|\mathbf{x}|^2\boldsymbol{\omega} - (\boldsymbol{\omega}\,.\,\mathbf{x})\mathbf{x}\big)\,\mathrm{d}V,$$

or, in suffix notation,

$$\begin{aligned} h_i &= \iiint\limits_V \rho(\mathbf{x})(x_k x_k \omega_i - \omega_j x_j x_i)\,\mathrm{d}V \\ &= \iiint\limits_V \rho(\mathbf{x})(x_k x_k \delta_{ij} - x_j x_i)\omega_j\,\mathrm{d}V \\ &= I_{ij}\omega_j \end{aligned}$$

where

$$I_{ij} = \iiint\limits_V \rho(\mathbf{x})(x_k x_k \delta_{ij} - x_i x_j)\,\mathrm{d}V$$

is the *inertia tensor* of the rigid body. Note that the tensor I does not depend on $\boldsymbol{\omega}$, only on properties of the body itself; so it may be calculated once and for all for any given body. To see that it is indeed a tensor, note that both $\mathbf{h}$ and $\boldsymbol{\omega}$ are vectors, and apply arguments previously used for the conductivity tensor.

(iv) Susceptibility χ. If $\mathbf{M}$ is the magnetization (magnetic moment per unit volume) and $\mathbf{B}$ is the applied magnetic field, then for a simple medium we have $\mathbf{M} = \chi^{(m)}\mathbf{B}$ where $\chi^{(m)}$ is the magnetic susceptibility. This generalises to $M_i = \chi^{(m)}_{ij} B_j$ where $\chi^{(m)}_{ij}$ is the *magnetic susceptibility tensor*. Similarly for polarization density $\mathbf{P}$ in a dielectric: $P_i = \chi^{(e)}_{ij} E_j$ where $\mathbf{E}$ is the electric field and $\chi^{(e)}_{ij}$ is the *electric susceptibility tensor*.

(v) The Kronecker delta itself. We have defined δ_{ij} without reference to frame; i.e., its components are by definition the same in all frames ($\delta'_{ij} \equiv \delta_{ij}$). Surprisingly, then, we can show that it is a tensor:

$$l_{ip} l_{jq} \delta_{pq} = l_{ip} l_{jp} = \delta_{ij} = \delta'_{ij}$$

(from §3.3), which is exactly the right transformation law. We can also show that ϵ_{ijk} is a tensor of rank 3.

Both δ_{ij} and ϵ_{ijk} are *isotropic tensors*: that is, their components are the same in all frames.

(vi) Stress and strain tensors. In an elastic body, stresses (forces) produce displacements of small volume elements within the body. Let this displacement at a location $\mathbf{x}$ be $\mathbf{u}$; then the *strain tensor* is defined to be

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

The stress tensor p_{ij} is defined as the j^{th} component of the forces within the body acting on an imaginary plane perpendicular to the i^{th} axis. Hooke's law for simple media says that stress $\propto$ strain. We can now generalise this to the tensor formulation

$$p_{ij} = k_{ijkl} e_{kl}$$

where k_{ijkl} is a fourth rank tensor, which expresses the linear (but possibly anisotropic) relationship between p and e.

3.5 Properties of Tensors

Linear Combination of Tensors

If A_{ij} and B_{ij} are second rank tensors, and α, β are scalars, then $T_{ij} = \alpha A_{ij} + \beta B_{ij}$ is a tensor.

Proof:

$$\begin{aligned} T'_{ij} &= \alpha' A'_{ij} + \beta' B'_{ij} \\ &= \alpha l_{ip} l_{jq} A_{pq} + \beta l_{ip} l_{jq} B_{pq} \\ &= l_{ip} l_{jq} (\alpha A_{pq} + \beta B_{pq}) \\ &= l_{ip} l_{jq} T_{pq} \end{aligned}$$

as required.

This result clearly extends to tensors of rank n.

Contraction (also known as the Inner Product)

If T_{ij} is a tensor then T_{ii} is a scalar. Proof:

$$T'_{ii} = l_{ip} l_{iq} T_{pq} = \delta_{pq} T_{pq} = T_{pp} = T_{ii},$$

so T_{ii} has the same value in all frames as required.

We can extend this result: if $T_{ijk\ldots lmn\ldots}$ is a tensor of rank n then $S_{jk\ldots mn\ldots} = T_{ijk\ldots imn\ldots}$ is a tensor of rank $n-2$. Proof:

$$\begin{aligned} S'_{jk\ldots mn\ldots} &= T'_{ijk\ldots imn\ldots} \\ &= l_{ip}l_{jq}l_{kr}\ldots l_{i\alpha}l_{m\beta}l_{n\gamma}\ldots T_{pqr\ldots\alpha\beta\gamma\ldots} \\ &= (l_{ip}l_{i\alpha})l_{jq}l_{kr}\ldots l_{m\beta}l_{n\gamma}\ldots T_{pqr\ldots\alpha\beta\gamma\ldots} \\ &= \delta_{p\alpha}l_{jq}l_{kr}\ldots l_{m\beta}l_{n\gamma}\ldots T_{pqr\ldots\alpha\beta\gamma\ldots} \\ &= l_{jq}l_{kr}\ldots l_{m\beta}l_{n\gamma}\ldots S_{qr\ldots\beta\gamma\ldots}. \end{aligned}$$

Outer Product

If $\mathbf{a}$ and $\mathbf{b}$ are vectors then the outer product T_{ij} defined by $T_{ij} = a_i b_j$ is a tensor of rank two. Proof:

$$T'_{ij} = a'_i b'_j = l_{ip}a_p l_{jq}b_q = l_{ip}l_{jq}a_p b_q = l_{ip}l_{jq}T_{pq}$$

as required.

Similarly (left as an exercise for the reader) we can show that if $A_{ijk\ldots}$ is a tensor of rank m and $B_{lmn\ldots}$ is a tensor of rank n, then $T_{ijk\ldots lmn\ldots} = A_{ijk\ldots}B_{lmn\ldots}$ is a tensor of rank $m+n$.

Example: if $\mathbf{a}$ and $\mathbf{b}$ are vectors then $\mathbf{a}\,.\,\mathbf{b}$ is a scalar. Proof: $T_{ij} = a_i b_j$, being an outer product of two vectors, is a tensor of rank two. Then $T_{ii} = a_i b_i$, being a contraction of a tensor, is a scalar, as required. Note that $|\mathbf{a}|^2 = \mathbf{a}\,.\,\mathbf{a}$ and $|\mathbf{b}|^2$ are also scalars; hence $\mathbf{a}\,.\,\mathbf{b}/|\mathbf{a}|\,|\mathbf{b}| = \cos\theta$ is a scalar, so that the angle between vectors is unaffected by a change of frame.

3.6 Symmetric and Anti-Symmetric Tensors

A tensor $T_{ijk\ldots}$ is said to be symmetric in a pair of indices (say i, j) if

$$T_{ijk\ldots} = T_{jik\ldots}$$

or anti-symmetric in i, j if

$$T_{ijk\ldots} = -T_{jik\ldots}.$$

For a second rank tensor we need not specify the indices as there are only two to choose from! For example, δ_{ij} is symmetric; ϵ_{ijk} is anti-symmetric in any pair of indices.

Note: if A_{ij} is a symmetric second rank tensor then the matrix corresponding to A is symmetric, i.e. $A = A^T$. Similarly for an anti-symmetric tensor.

Suppose that S_{ij} is a symmetric tensor and A_{ij} an anti-symmetric tensor. Then $S_{ij}A_{ij} = 0$. Proof:

$$\begin{aligned} S_{ij}A_{ij} = -S_{ij}A_{ji} &= -S_{ji}A_{ji} \\ &= -S_{ij}A_{ij} \qquad \text{(swapping dummy } i \text{ and } j\text{)} \\ \implies \qquad 2S_{ij}A_{ij} &= 0, \end{aligned}$$

as required. Try to work out also how to see this "by inspection", by considering appropriate pairs of components.

Example: for any vector $\mathbf{a}$, $\mathbf{a} \times \mathbf{a} = \mathbf{0}$ because

$$[\mathbf{a} \times \mathbf{a}]_i = \epsilon_{ijk} a_j a_k$$

and ϵ_{ijk} is anti-symmetric in j, k whilst $a_j a_k$ is symmetric.

The properties of symmetry and anti-symmetry are invariant under a change of frame: that is, they are truly tensor properties. For example, suppose that T_{ij} is symmetric. Then

$$\begin{aligned} T'_{ij} &= l_{ip} l_{jq} T_{pq} \\ &= l_{ip} l_{jq} T_{qp} \\ &= l_{jq} l_{ip} T_{qp} = T'_{ji}, \end{aligned}$$

so that T'_{ij} is also symmetric.

(Alternative, and simpler, proof for second rank tensors:

$$T' = LTL^T \qquad \implies \qquad T'^T = (LTL^T)^T = LT^T L^T = LTL^T = T'$$

using $T^T = T$.)

Symmetry and anti-symmetry occur frequently in practical applications. For example, the strain tensor $e_{ij} = \frac{1}{2}(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$ is clearly symmetric. In most situations the stress tensor is also symmetric; but in some circumstances (for instance in crystallography or geodynamics) it is forced to be anti-symmetric while the strain remains symmetric. Inertia tensors are always symmetric; conductivity and susceptibility tensors usually are.

Decomposition into Symmetric and Anti-Symmetric Parts

Any second rank tensor T_{ij} can be uniquely expressed as the sum of a symmetric and an anti-symmetric tensor; for

$$\boxed{T_{ij} = S_{ij} + A_{ij}}$$

where

$$\boxed{S_{ij} = \tfrac{1}{2}(T_{ij} + T_{ji}), \qquad A_{ij} = \tfrac{1}{2}(T_{ij} - T_{ji})}$$

are symmetric and anti-symmetric respectively. Exercise: prove that S and A are tensors.

Furthermore, any anti-symmetric tensor A_{ij} can be expressed in terms of a vector $\boldsymbol{\omega}$ (sometimes known as the *dual vector*) such that

$$\boxed{A_{ij} = \epsilon_{ijk}\omega_k.}$$

Proof: define $\boldsymbol{\omega}$ by

$$\omega_k = \tfrac{1}{2}\epsilon_{klm}A_{lm}.$$

Then

$$\begin{aligned}
\epsilon_{ijk}\omega_k &= \tfrac{1}{2}\epsilon_{ijk}\epsilon_{klm}A_{lm} \\
&= \tfrac{1}{2}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})A_{lm} \\
&= \tfrac{1}{2}(A_{ij} - A_{ji}) = A_{ij}
\end{aligned}$$

as required. $\boldsymbol{\omega}$ is a vector as it is a contraction of two tensors.

This definition of $\boldsymbol{\omega}$ actually corresponds to setting

$$A = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

Example: suppose that two symmetric second rank tensors R_{ij} and S_{ij} are linearly related. Then there must be a relationship between them of the form $R_{ij} = c_{ijkl}S_{kl}$. It is clear that c_{ijkl} must be symmetric in i, j (for otherwise, R_{ij} would not be). It is not necessarily the case that it must also be symmetric in k, l, but without loss of generality we may assume that it is, by the following argument. Decompose c_{ijkl} into a part $c^{(\mathrm{s})}_{ijkl}$ which is symmetric in k, l and a part $c^{(\mathrm{a})}_{ijkl}$ which is anti-symmetric. Then

$$R_{ij} = c^{(\mathrm{s})}_{ijkl}S_{kl} + c^{(\mathrm{a})}_{ijkl}S_{kl} = c^{(\mathrm{s})}_{ijkl}S_{kl}$$

because the second term is the contraction of an anti-symmetric tensor with a symmetric one, which we showed was zero above. Hence we can ignore any anti-symmetric part of c_{ijkl}.

3.7 Diagonalization of Symmetric Second Rank Tensors

Suppose T_{ij} is a symmetric second rank tensor. We shall show that there exists a frame such that, if we transform T to that frame, it has components given by

$$T' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

This process is known as diagonalization. The values λ_1, λ_2 and λ_3 are known as the *principal values* of T, and the Cartesian coordinate axes of the corresponding frame are known as the *principal axes*. We will see that in fact the principal values are just the eigenvalues of the matrix corresponding to T, and the principal axes are the eigenvectors.

Because T is symmetric, we know that there are 3 real eigenvalues and that we can find 3 corresponding eigenvectors which are orthogonal and of unit length. Let λ_1, λ_2, λ_3 be the eigenvalues and $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ be the eigenvectors (arranged as a right-handed set of orthonormal vectors). Change frame to one in which the coordinate axes are aligned with $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. What is T'?

Recall that $L^T = \left(\mathbf{e}'_1 \mid \mathbf{e}'_2 \mid \mathbf{e}'_3 \right)$; i.e., the three columns of L^T are the vectors $\mathbf{e}'_1$, $\mathbf{e}'_2$ and $\mathbf{e}'_3$ (measured relative to the first frame). Hence in matrix notation,

$$\begin{aligned} TL^T &= T\left(\mathbf{e}'_1 \mid \mathbf{e}'_2 \mid \mathbf{e}'_3 \right) \\ &= \left(\lambda_1\mathbf{e}'_1 \mid \lambda_2\mathbf{e}'_2 \mid \lambda_3\mathbf{e}'_3 \right). \end{aligned}$$

So

$$\begin{aligned} T' = LTL^T &= \left(\begin{array}{c} \mathbf{e}'^T_1 \\ \hline \mathbf{e}'^T_2 \\ \hline \mathbf{e}'^T_3 \end{array} \right) \left(\begin{array}{c|c|c} \lambda_1\mathbf{e}'_1 & \lambda_2\mathbf{e}'_2 & \lambda_3\mathbf{e}'_3 \end{array} \right) \\ &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \end{aligned}$$

because, for example, the top LHS entry is given by $\mathbf{e}'_1 \cdot \lambda_1\mathbf{e}'_1$, and the top RHS entry is $\mathbf{e}'_1 \cdot \lambda_3\mathbf{e}'_3$.

There is another way of seeing that $T' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$. The equation $T\mathbf{e}'_1 = \lambda_1\mathbf{e}'_1$ is true in any frame (because T is a tensor, $\mathbf{e}'_1$ a vector and λ_1 a scalar). In particular it is true in the frame with $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$

as coordinate axes. But, measured in this frame, $\mathbf{e}'_1$ is just $(1, 0, 0)^T$, and T has components T'; so

$$T' \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix}$$

which shows that the first column of T' is $(\lambda_1, 0, 0)^T$. Similarly for the other columns.

Note: the three principal values are invariants of T. That is, whatever frame we start from, when we diagonalize T we will obtain the same values of λ. The eigenvalues are properties of the tensor, not of the coordinate system.

3.8 Isotropic Tensors

An *isotropic tensor* is one whose components are the same in all frames, i.e.,

$$\boxed{T'_{ijk\ldots} = T_{ijk\ldots}.}$$

We can classify isotropic tensors up to rank four as follows:

Rank 0: All scalars are isotropic, since the tensor transformation law states that $T' = T$ for tensors of rank zero.

Rank 1: There are no non-zero isotropic vectors.

Rank 2: The most general isotropic second rank tensor is $\boxed{\lambda\delta_{ij}}$ where λ is any scalar, as proved below.

Rank 3: The most general isotropic third rank tensor is $\boxed{\lambda\epsilon_{ijk}.}$

Rank 4: The most general isotropic fourth rank tensor is

$$\lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl} + \nu\delta_{il}\delta_{jk}$$

where λ, μ, ν are scalars.

What is the physical significance of an isotropic tensor? Consider the conductivity tensor σ_{ij} in an isotropic medium. As the medium is the same in all directions, we expect that σ_{ij} will be isotropic too. Hence $\sigma_{ij} = \lambda\delta_{ij}$ and

$$J_i = \sigma_{ij}E_j = \lambda\delta_{ij}E_j = \lambda E_i,$$

i.e., $\mathbf{J} = \lambda\mathbf{E}$. So we recover the "simple version" of the conductivity law, as we might expect.

Isotropic Second Rank Tensors

Consider a general tensor T of rank two, with components T_{ij} with respect to some set of axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Suppose that T is isotropic. Its components should then be unaltered by a rotation of 90° about the 3-axis, i.e., with respect to new axes

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

The matrix of this rotation is

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the matrix formulation of the transformation law for tensors, we see that

$$\begin{aligned} \begin{pmatrix} T'_{11} & T'_{12} & T'_{13} \\ T'_{21} & T'_{22} & T'_{23} \\ T'_{31} & T'_{32} & T'_{33} \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} T_{22} & -T_{21} & T_{23} \\ -T_{12} & T_{11} & -T_{13} \\ T_{32} & -T_{31} & T_{33} \end{pmatrix}. \end{aligned}$$

But, because T is isotropic, $T'_{ij} = T_{ij}$. Hence, comparing matrix entries, we have:

$$T_{11} = T_{22};$$

$$T_{13} = T_{23} = -T_{13} \qquad \text{so that} \qquad T_{13} = T_{23} = 0;$$

$$T_{31} = T_{32} = -T_{31} \qquad \text{so that} \qquad T_{31} = T_{32} = 0.$$

Similarly, considering a rotation of 90° about the 2-axis, we find that $T_{11} = T_{33}$ and that $T_{12} = T_{32} = 0$, $T_{21} = T_{23} = 0$. Therefore *all* off-diagonal elements of T are zero, and all diagonal elements are equal, say λ. Thus

$$T = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

or in suffix notation, $T_{ij} = \lambda\delta_{ij}$.

In summary, we have shown that *any* isotropic second rank tensor must be equal to $\lambda\delta_{ij}$ for some scalar λ.

3.9 Tensor Differential Operators

A *tensor field* is a tensor which depends on the location $\mathbf{x}$. For example:

(i) Temperature is a scalar field (a tensor field of rank zero), because $T = T(\mathbf{x})$.

(ii) Any vector field $\mathbf{F}(\mathbf{x})$, such as a gravitational force field, is a tensor field of rank one. In particular, $\mathbf{x}$ is itself a vector field, because it is a vector function of position!

(iii) In a conducting material where the conductivity varies with location, we have $\sigma_{ij} = \sigma_{ij}(\mathbf{x})$, a tensor field of rank two.

We are interested here in calculating the derivatives of tensor fields; we start with scalars and vectors.

We can rewrite the definitions of grad, div, and curl using suffix notation.

$$\textbf{Grad:}\qquad [\nabla\Phi]_i = \frac{\partial\Phi}{\partial x_i}$$

$$\textbf{Div:}\qquad \nabla\,.\,\mathbf{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} = \frac{\partial F_i}{\partial x_i}$$

$$\textbf{Curl:}\qquad [\nabla\times\mathbf{F}]_i = \epsilon_{ijk}\frac{\partial F_k}{\partial x_j}$$

There is another notation worth knowing: if $\mathbf{u}$, $\mathbf{v}$ are vectors then we define the vector

$$(\mathbf{u}\,.\,\nabla)\mathbf{v} = \left(u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2} + u_3\frac{\partial}{\partial x_3}\right)\mathbf{v}.$$

In suffix notation,

$$\boxed{[(\mathbf{u}\,.\,\nabla)\mathbf{v}]_i = u_j\frac{\partial v_i}{\partial x_j}.}$$

Laplace's equation $\nabla^2\Phi = 0$ becomes

$$\frac{\partial^2\Phi}{\partial x_i\partial x_i} = 0$$

in suffix notation. Similarly,

$$[\nabla^2\mathbf{F}]_i = \frac{\partial^2 F_i}{\partial x_j\partial x_j}$$

(note that we only use Cartesian coordinates here).

We sometimes find it useful to use the *differential operator* ∂_i defined by

$$\partial_i = \frac{\partial}{\partial x_i}.$$

Then

$$\boxed{[\nabla\Phi]_i = \partial_i\Phi; \qquad \nabla\,\boldsymbol{.}\,\mathbf{F} = \partial_i F_i; \qquad [\nabla\times\mathbf{F}]_i = \epsilon_{ijk}\partial_j F_k.}$$

It turns out that ∂_i is in fact a tensor of rank one. We know that $x_j = l_{ij}x'_i$ (from $\mathbf{x} = L^T\mathbf{x}'$) so that

$$\frac{\partial x_j}{\partial x'_i} = \frac{\partial}{\partial x'_i}(l_{kj}x'_k) = l_{kj}\frac{\partial x'_k}{\partial x'_i} = l_{kj}\delta_{ik} = l_{ij}.$$

(This looks obvious but has to be proved very carefully!) Now let T be some quantity (perhaps a scalar or a tensor of some rank). Then

$$\begin{aligned}\partial'_i T = \frac{\partial T}{\partial x'_i} &= \frac{\partial T}{\partial x_1}\frac{\partial x_1}{\partial x'_i} + \frac{\partial T}{\partial x_2}\frac{\partial x_2}{\partial x'_i} + \frac{\partial T}{\partial x_3}\frac{\partial x_3}{\partial x'_i}\\ &= \frac{\partial T}{\partial x_j}\frac{\partial x_j}{\partial x'_i}\\ &= l_{ij}\frac{\partial T}{\partial x_j} = l_{ij}\partial_j T.\end{aligned}$$

This is true for any quantity T, so

$$\partial'_i = l_{ij}\partial_j,$$

i.e., ∂_i transforms like a vector, and is hence a tensor of rank one.

This result allows us to prove that $\nabla\Phi$, $\nabla\,\boldsymbol{.}\,\mathbf{F}$ and $\nabla\times\mathbf{F}$ are scalars or vectors (as appropriate). For example, to show that if $\mathbf{F}$ is a vector field then $\nabla\times\mathbf{F}$ is a vector field:

$$\begin{aligned}[\nabla\times\mathbf{F}]'_i &= \epsilon'_{ijk}\partial'_j F'_k\\ &= l_{ip}l_{jq}l_{kr}\epsilon_{pqr}l_{js}\partial_s l_{kt}F_t\\ &\qquad\qquad [\epsilon,\ \partial \text{ and } \mathbf{F} \text{ are tensors}]\\ &= l_{ip}(l_{jq}l_{js})(l_{kr}l_{kt})\epsilon_{pqr}\partial_s F_t\\ &= l_{ip}\delta_{qs}\delta_{rt}\epsilon_{pqr}\partial_s F_t\\ &= l_{ip}\epsilon_{pqr}\partial_q F_r\\ &= l_{ip}[\nabla\times\mathbf{F}]_p,\end{aligned}$$

as required.

Alternatively, we can just state that $\nabla\times\mathbf{F}$ is a contraction of the tensor outer product $T_{ijklm} = \epsilon_{ijk}\partial_l F_m$ (because $[\nabla\times\mathbf{F}]_i = T_{ijkjk}$).

As an example of a tensor field of rank three, consider the derivative of the conductivity tensor, $\partial_i\sigma_{jk}$. This cannot be written using ∇.