Here are some notes on vector and dyadic tensor notation similar to what I will be using in class, with just a couple of changes in notation.

Notes = Dowling

\[ \land = \times \quad \text{British notation for Cross-Product vs. American (Dowling's)} \]

\[ \vec{A} = \vec{A} \quad \text{Notation for Vector vs. Standard (Dowling's).} \]

\[ \overline{B} = \overline{B} \quad \text{Notation for Dyadic Tensor vs. Standard (Dowling's).} \]

With these notes you should be able to make sense of the expressions in the vector identities pages below that involve the dyadic tensor \( T \).

(Notes courtesy of Peter Littlewood, University of Cambridge.)

The vector identity pages are from the NRL Plasma Formulary – order your very own online or download the entire thing at:


--JPD
VECTOR IDENTITIES

Notation: $f$, $g$, are scalars; $\mathbf{A}$, $\mathbf{B}$, etc., are vectors; $\mathbf{T}$ is a tensor; $I$ is the unit dyad.

(1) $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{A} \times \mathbf{B} \cdot \mathbf{C} = \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{B} \times \mathbf{C} \cdot \mathbf{A} = \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} = \mathbf{C} \times \mathbf{A} \cdot \mathbf{B}$

(2) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{C} \times \mathbf{B}) \times \mathbf{A} = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C}$

(3) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = 0$

(4) $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C}) (\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D}) (\mathbf{B} \cdot \mathbf{C})$

(5) $(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \times \mathbf{B} \cdot \mathbf{D}) \mathbf{C} - (\mathbf{A} \times \mathbf{B} \cdot \mathbf{C}) \mathbf{D}$

(6) $\nabla (fg) = \nabla (gf) = f \nabla g + g \nabla f$

(7) $\nabla \cdot (f \mathbf{A}) = f \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f$

(8) $\nabla \times (f \mathbf{A}) = f \nabla \times \mathbf{A} + \nabla f \times \mathbf{A}$

(9) $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$

(10) $\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} (\nabla \cdot \mathbf{B}) - \mathbf{B} (\nabla \cdot \mathbf{A}) + (\mathbf{B} \cdot \nabla) \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B}$

(11) $\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B}$

(12) $\nabla (\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A}$

(13) $\nabla^2 f = \nabla \cdot \nabla f$

(14) $\nabla^2 \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \nabla \times \nabla \times \mathbf{A}$

(15) $\nabla \times \nabla f = 0$

(16) $\nabla \cdot \nabla \times \mathbf{A} = 0$

If $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are orthonormal unit vectors, a second-order tensor $\mathbf{T}$ can be written in the dyadic form

(17) $\mathbf{T} = \sum_{i,j} T_{ij} \mathbf{e}_i \mathbf{e}_j$

In cartesian coordinates the divergence of a tensor is a vector with components

(18) $(\nabla \cdot \mathbf{T})_i = \sum_j (\partial T_{ij} / \partial x_j)$

[This definition is required for consistency with Eq. (29)]. In general

(19) $\nabla \cdot (\mathbf{A} \mathbf{B}) = (\nabla \cdot \mathbf{A}) \mathbf{B} + (\mathbf{A} \cdot \nabla) \mathbf{B}$

(20) $\nabla \cdot (f \mathbf{T}) = \nabla f \cdot \mathbf{T} + f \nabla \cdot \mathbf{T}$
Let $\mathbf{r} = ix + jy + kz$ be the radius vector of magnitude $r$, from the origin to the point $x, y, z$. Then

\begin{align*}
(21) \nabla \cdot \mathbf{r} &= 3 \\
(22) \nabla \times \mathbf{r} &= 0 \\
(23) \nabla r &= r/r \\
(24) \nabla(1/r) &= -r/r^3 \\
(25) \nabla \cdot (r/r^3) &= 4\pi\delta(\mathbf{r}) \\
(26) \nabla \mathbf{r} &= \mathbf{l}
\end{align*}

If $V$ is a volume enclosed by a surface $S$ and $d\mathbf{S} = \mathbf{n} \, dS$, where $\mathbf{n}$ is the unit normal outward from $V$,

\begin{align*}
(27) \int_V dV \nabla f &= \int_S d\mathbf{S} f \\
(28) \int_V dV \nabla \cdot \mathbf{A} &= \int_S d\mathbf{S} \cdot \mathbf{A} \\
(29) \int_V dV \nabla \cdot \mathbf{T} &= \int_S d\mathbf{S} \cdot \mathbf{T} \\
(30) \int_V dV \nabla \times \mathbf{A} &= \int_S d\mathbf{S} \times \mathbf{A} \\
(31) \int_V dV (f \nabla^2 g - g \nabla^2 f) &= \int_S d\mathbf{S} \cdot (f \nabla g - g \nabla f) \\
(32) \int_V dV (\mathbf{A} \cdot \nabla \times \nabla \times \mathbf{B} - \mathbf{B} \cdot \nabla \times \nabla \times \mathbf{A}) \\
&= \int_S d\mathbf{S} \cdot (\mathbf{B} \times \nabla \times \mathbf{A} - \mathbf{A} \times \nabla \times \mathbf{B})
\end{align*}

If $S$ is an open surface bounded by the contour $C$, of which the line element is $dl$,

\begin{align*}
(33) \int_S d\mathbf{S} \times \nabla f &= \oint_C dl f
\end{align*}
(34) \[ \int_S dS \cdot \nabla \times A = \oint_C dl \cdot A \]

(35) \[ \int_S (dS \times \nabla) \times A = \oint_C dI \times A \]

(36) \[ \int_S dS \cdot (\nabla f \times \nabla g) = \oint_C f dg = -\oint_C g df \]
V. Introduction to Tensors (3.2.99)

Commentary: This is the first of two sheets on tensors (the second will come in class VIII). Tensors have many applications in theoretical physics, not only in polarization problems like the ones described below, but in dynamics, elasticity, fluid mechanics, quantum theory, etc.. The mathematical apparatus for dealing with tensor problems may not yet be familiar, so this sheet contains quite a high proportion of formal material.

Section A introduces the concept of a tensor (in this case, a linear operator that transforms one vector into another) in the physical context of studying the polarisation of an aspherical molecule.

Section B summarizes various mathematical and notational ideas that are useful in physics; these include suffix notation, and concepts such as the outer product of two vectors, and projection operators. (These are mathematical objects which take the projection or component of a vector along a particular direction.)

Section C returns to a physical problem (also about polarization) and leads to a discussion of principal axes, eigenvectors and eigenvalues of tensor quantities.

Section D introduces the properties of the antisymmetric third-rank tensor $\epsilon_{ijk}$ which can be used to derive many results in vector calculus much more rapidly than by other methods. This ground will be covered again in sheet VIII so it doesn’t matter if you run out of time.
A. Introduction

We often find vectors that are linearly related to other vectors, for instance the polarisation $\mathbf{P}$ induced in a medium by applying a field $\mathbf{E}$

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E}$$  \hspace{1cm} (1)

Here $\mathbf{P}$ and $\mathbf{E}$ are parallel. Their cartesian components are related to each other by the same constant of proportionality $\epsilon_0 \chi$ for each component. [Here we have denoted vectors by underlinings as one would do if writing this out by hand. Usually books denote vectors by bold type; in the rest of this sheet we treat the two notations as completely interchangeable.]

It is possible to imagine situations where the medium is anisotropic and the response factor $\epsilon_0 \chi$ in (1) is different for components in different directions. The two vectors $\mathbf{P}$ and $\mathbf{E}$ are no longer necessarily parallel and what connects them is called a tensor. We will use simple model of molecular polarisability to show the underlying physics and introduce the ideas of tensors.

A1. An Electrostatics Problem. The diacetylene molecule, H-C≡C-C≡C-H is highly polarisable along its long axis, $p_\parallel = \beta_\parallel E_\parallel$ being the induced dipole for a field $E_\parallel$ applied along the molecule. The response to perpendicular fields is $p_\perp = \beta_\perp E_\perp$ where $\beta_\perp \ll \beta_\parallel$. Let the cylindrical molecule point in the $(1, 1, 0)$ direction and apply a field in the $x$ direction: $\mathbf{E} = E_x \hat{x}$

(i) Prove:

$$p_x = (\beta_\parallel + \beta_\perp)E_x/2$$  \hspace{1cm} (2)

$$p_y = (\beta_\parallel - \beta_\perp)E_x/2$$  \hspace{1cm} (3)

$$p_z = 0$$  \hspace{1cm} (4)
Hence one sees that $p$ and $\mathbf{E}$ are not parallel and in general we must write

$$p = \alpha \cdot \mathbf{E}$$

where $\alpha$ is the polarisability tensor. Tensors are written with a double underline (as matrices often are) or in books sometimes as “doubly bold” (bold sans serif typeface) characters. In component form (5) is

$$p_x = \alpha_{xx}E_x + \alpha_{xy}E_y + \alpha_{xz}E_z$$

$$p_y = \alpha_{yx}E_x + \alpha_{yy}E_y + \alpha_{yz}E_z$$

$$p_z = \alpha_{zx}E_x + \alpha_{zy}E_y + \alpha_{zz}E_z$$

(ii) What are the coefficients $\alpha_{ij}$ explicitly for the above example? Write $\alpha$ as a matrix of its components. Note that, although the tensor $\alpha$ is the same physical entity in all coordinate systems (it relates $p$ with $\mathbf{E}$ in whatever basis these are expressed), its components $\alpha_{ij}$ depend on the coordinate system used.

**Remark:** If many molecules are present with the same orientation, $\mathbf{P}$ is $\rho \alpha$ where there are $\rho$ molecules per unit volume. A (hypothetical) diacetylene solid might consist of such rods arranged in an array, all directed along $(1, 1, 0)$. Then the macroscopic $\chi$ (to replace $\chi$ in (1)) would (neglecting internal field corrections) be $\rho \alpha$.

**B. Miscellaneous Ideas and Notations**

**B1. Suffix notation:** Commonly, suffices such as $i$ and $j$ are dummy symbols and when repeated in an expression they are to be summed over; this is the **Einstein Convention**. For instance (6) to (8) can be succinctly written as

$$p_i = \alpha_{ij}E_j$$
which is Eq. (5) rewritten in suffix notation. (Note the order of the indices on the right.) Thus we have written the vector $\mathbf{p}$ simply as $p_i$ and it will be clear from context that a vector is intended and not simply one of its components. Likewise $\mathbf{a}$ can be denoted $a_{ij}$.

In suffix notation the dot product (also called scalar product, or *inner product*) of two vectors $\mathbf{a} \cdot \mathbf{b}$ is written as $a_i b_i$. The operation of a tensor on a vector also involves an inner product (sum over repeated indices) as described above. The *unit tensor* $\delta_{ij}$ by definition transforms any vector into itself: $\delta_{ij} v_j = v_i$.

**B2. The outer product:** Consider the object $Q_{ij}$ defined as

$$Q_{ij} \equiv a_i b_j$$

where $\mathbf{a}$ and $\mathbf{b}$ are vectors.

(i) Show using suffix notation that if $\mathbf{p}$ is a vector, $Q \cdot \mathbf{p}$ is also a vector; find its magnitude and direction. [Note that $Q \cdot \mathbf{p} \equiv Q_{ij} p_j$.] We see that $Q_{ij}$ is a linear operator that transforms one vector $\mathbf{p}$ into another; it is therefore a tensor. This defines the *outer product* or *dyadic product* of two vectors.

In non-suffix notation (sometimes called dyadic notation), Eq. (9) is written as

$$\mathbf{Q} = \mathbf{ab}$$

or $\mathbf{Q} = \mathbf{a} \mathbf{b}$, where there is no dot between the two vectors. In this notation we have $(\mathbf{a} \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} (\mathbf{b} \cdot \mathbf{c})$: prove this using suffix notation if you did not do so already above. This notation is used in various areas of mathematics and physics, but in general suffix notation is more versatile and can always be used instead.

(ii) We can write the tensor $\underline{\mathbf{a}}$ in $\mathbf{Q}$. A1 in the following form:

$$\underline{\mathbf{a}} = \alpha_{xx} \hat{x} \hat{x} + \alpha_{xy} \hat{x} \hat{y} + \alpha_{xz} \hat{x} \hat{z} + 6 \text{ other terms}$$

Check that (11) indeed agrees with (6) to (8) by finding the result of applying a field in the $x$-direction, $\mathbf{E} = E_x \hat{x}$:

$$\mathbf{p} = \underline{\mathbf{a}} \cdot \mathbf{E} = \underline{\mathbf{a}} \cdot \hat{x} E_x = (\alpha_{xx}, \alpha_{yx}, \alpha_{zx}) E_x.$$
(iii) **Projection operators:** In dyadic notation, the identity tensor $\delta_{ij}$ is often written as the “unit operator” $\mathbb{1}$. Consider the action on a vector $\mathbf{v}$ of the operators $\hat{u}\hat{u}$ and $\mathbb{1} - \hat{u}\hat{u}$, where $\hat{u}$ is a unit vector. Show that these perform the job of resolving the vector $\mathbf{v}$ into vectors $\mathbf{v}_\parallel$ and $\mathbf{v}_\perp$, respectively parallel and perpendicular to $\hat{u}$. Prove that these “projection operators” are *idempotent*. [The definition of an idempotent operator $G$ is that $GG = G$.] Show that the polarizability tensor in problem A1 can be written

$$\alpha = \beta_\parallel \hat{u}\hat{u} + \beta_\perp (\mathbb{1} - \hat{u}\hat{u})$$

where $\hat{u} = (1,1,0)/\sqrt{2}$.

\[ C. \text{ Use of Principal Axes.} \]

**C1.** A fixed object is placed in a uniform electric field of 1 kV/m. When the field is in the $x$ direction, it acquires a dipole moment with components 4, 2, and $1 \times 10^{-12}$ C m in the $x$, $y$, $z$ directions. For the same field strengths in the $y$, $z$ directions, the dipole moments are $(2,4,1) \times 10^{-12}$ and $(1,1,4) \times 10^{-12}$ C m respectively.

(i) Write down the components of the tensor that describes the polarisability of the object (using the $x,y,z$ coordinate frame).

(ii) What is the torque $\mathbf{T} = \mathbf{p} \wedge \mathbf{E}$ on the object when it is placed in an electric field of $\sqrt{3}$ kV/m in the $(1,1,1)$ direction? [Ans: $(1,-1,0) \times 10^{-9}$ Nm.]

(iii) Suppose there are field directions $\hat{u}, \hat{v}, \hat{w}$ for which there is no torque on the object; what relation do $\mathbf{E}$ and $\mathbf{p}$ then have? Show that $\hat{u}, \hat{v}, \hat{w}$ are the eigenvectors of $\underline{\underline{\alpha}}$.

(iv) Show that, if we use as a coordinate system the unit vectors $\hat{u}, \hat{v}, \hat{w}$ there are no off-diagonal elements of the tensor $\underline{\underline{\alpha}}$ when it is written as a matrix of components. The three diagonal elements, $\lambda_1, \lambda_2$ and $\lambda_3$ are called the *principal polarisabilities* of the object and are the eigenvalues of $\underline{\underline{\alpha}}$. For this problem, one eigenvalue is $\lambda = 2 \times 10^{-15}$. find the
other two. What are the vectors \( \mathbf{u}, \mathbf{v}, \mathbf{w} \) in terms of \( \mathbf{x}, \mathbf{y}, \mathbf{z} \)?

(v) In this example, \( \underline{\alpha} \) is symmetric and hence has real eigenvalues and orthogonal eigenvectors. An energy argument shows this to be true for the polarizability of any object; the \( \mathbf{u}, \mathbf{v}, \mathbf{w} \) coordinate system can hence always be found and is called the principal frame - often a very convenient frame to think about the physics in. The same applies to any problem involving a symmetric tensor.

(vi) In a general coordinate system, the decomposition of \( \underline{\alpha} \) into principal directions \( \mathbf{u}, \mathbf{v}, \mathbf{w} \) can be accomplished using projection operators:

\[
\underline{\alpha} = \lambda_1 \mathbf{u} \mathbf{u} + \lambda_2 \mathbf{v} \mathbf{v} + \lambda_3 \mathbf{w} \mathbf{w}
\]

This is a relation between tensors, so if it holds in one co-ordinate system it also holds in any other. Prove it for a convenient choice of co-ordinates system (and hence generally).

(vii) Eq.(12) expresses in a very direct way the operation of a symmetric tensor \( \underline{\alpha} \) on a vector \( \mathbf{b} \). To find \( \underline{\alpha} \cdot \mathbf{b} \), the recipe is: (1) resolve \( \mathbf{b} \) into components along the principal axes; (2) multiply each component by the corresponding eigenvalue; (3) add to form the resultant. Confirm from Eq.(12) that this is precisely what operating with \( \underline{\alpha} \) on \( \mathbf{b} \) does.

C2. Force on a Dipole.

(i) By using suffix notation, and the ordinary chain rule for derivatives, prove the vector relation

\[
\nabla \cdot (\psi \mathbf{A}) = \psi \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla \psi
\]

Derive a similar result for \( \nabla (\mathbf{A} \cdot \mathbf{B}) \). [It’s best to leave the answer in suffix notation.]

(ii) In a uniform electric field, \( \mathbf{E} \), a dipole \( \mathbf{p} \) experiences a couple \( \mathbf{p} \times \mathbf{E} \) but no net force. Show that in a non-uniform field, a dipole experiences a net force

\[
F_i = \frac{\partial E_i}{\partial x_j} p_j
\]

and explain why \( \frac{\partial E_i}{\partial x_j} \) is a tensor. [This is very general; the gradient of a vector is a second-rank tensor.]
(iii) A more general expression for the force on a permanent dipole, derived from the energy, is $F = \text{grad}(p \cdot E)$, or:

$$F_j = \frac{\partial}{\partial x_j} p_i E_i.$$ 

Show that this is equivalent to Eq(13) only in the absence of a time-varying magnetic field.

(iv) A speck of matter has no permanent dipole moment, but the field $E$ induces a moment $\mathbf{p}$ in it. Under what conditions can we write (a) $p_i = \alpha_{ij} E_j$; (b) $p_i = \alpha E_i$? Write down the force in each case. Under what conditions can this be written $F = \alpha \text{grad}(E^2/2)$?

### D. The antisymmetric three-tensor $\epsilon_{ijk}$.

A tensor with two indices such as $\delta_{ij}$ is called a tensor of second rank, and its linear operation on a vector gives another vector. A vector is technically a tensor of rank one; a scalar of rank zero. Tensors of higher rank can also be defined. For example a tensor of rank four, $Q_{ijkl}$ acts on a rank two tensor to give another rank two tensor, $B_{ij} = Q_{ijkl} A_{kl}$ (with summation over repeated indices). For tensors higher than rank two, dyadic notation becomes extremely cumbersome and suffix notation is essential.

**D1.** An important and useful tensor of rank three is the fully antisymmetric three tensor denoted $\epsilon_{ijk}$. This operates on either a second rank tensor or two vectors to give a vector. For $i \neq j \neq k$ it is defined as $\epsilon_{ijk} = 1$ if $ijk$ is a cyclic permutation of $x, y, z$ (e.g., $yzx$) and $= -1$ if the order is anticyclic (e.g., $xzy$). If any two of $ijk$ are the same then $\epsilon_{ijk} = 0$ (e.g., $\epsilon_{xxy} = 0$). The tensor $\epsilon_{ijk}$ is very useful for writing down a vector product in suffix notation. Show that

$$(\mathbf{A} \wedge \mathbf{B})_i = \epsilon_{ijk} A_j B_k \quad (14)$$

The following relation

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \quad (15)$$
is well worth remembering (you should come across this in maths lectures). It saves a great deal of time when manipulating vector identities such as

\[ A \wedge (B \wedge C) = \epsilon_{ijk} A_j \epsilon_{klm} B_l C_m \]

\[ = A_j B_l C_m (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \]

\[ = B_i A_j C_j - C_i A_j B_j = (A \cdot C)B - (A \cdot B)C \]

which is a familiar result normally derived by far more tedious means.

The result Eq.(15) is especially helpful when dealing with vector identities involving the \( \nabla \) operator. Because \( \epsilon_{ijk} \) is a constant tensor, it commutes with \( \nabla \), and expressions can be re-ordered accordingly. For example

\[ \nabla \cdot (E \wedge H) \equiv \nabla_i (\epsilon_{ijk} E_j H_k) \]

\[ = \epsilon_{ijk} \nabla_i (E_j H_k) = \epsilon_{ijk} [E_j \nabla_i H_k + H_k \nabla_i E_j] \]

\[ = -E_j \epsilon_{ijk} \nabla_i H_k + H_k \epsilon_{kij} \nabla_i E_j \]

\[ = -E \cdot \nabla \wedge H + H \cdot \nabla \wedge E \]

which is a result used in in connection with Poynting’s vector in electromagnetism.

(i) Make sure you understand each step of this derivation.

(ii) Show, using similar methods, that

\[ \nabla \wedge (\nabla \wedge A) = \nabla(\nabla \cdot A) - \nabla^2 A \]

\[ \nabla \wedge (\psi A) = \psi \nabla \wedge A - A \wedge \nabla \psi \]

These relations are useful in electromagnetism and other topics in IB, and in the parts II and III courses.
Commentary: This is the second sheet on tensors. It follows on from Sheet 5. Section D of that sheet (on uses of the antisymmetric 3-tensor, $\epsilon_{ijk}$) is reprinted here as an appendix, and you should complete it before starting the main part of this sheet, if you didn’t do so last time.

As on sheet 5 we use underlining and bold notation interchangeably. Suffix notation and the summation convention are also used. Carets (^) denote unit vectors.

In Sections A and B tensor ideas are used to work through some results in rigid body dynamics. You should have covered this material in IB dynamics courses, and the formal development given here is good practice and will help deepen your understanding of moments of inertia, etc., as well as of tensors.

Section C introduces the concepts of stress and strain tensors in a solid. The linear relation between these is governed by a fourth rank tensor, whose properties are investigated.

Section D contains problems on the transformation properties of tensors: in mathematical courses, you may have seen tensors defined in terms of these transformation properties.
A. Rigid Body Rotations

A rigid body rotates with angular velocity $\omega$ about an axis through its centre of mass at the origin, 0. The vector $\omega$ points along the axis of rotation and determines the velocity of rotation at $\mathbf{r}$ according to $\mathbf{v} = \omega \times \mathbf{r}$. [Make sure you can derive this if you are not sure where it comes from.]

(i) Consider the body as a collection of point masses $\delta m_\mathbf{r}$ with position vectors $\mathbf{r}$ relative to the centre of mass 0.

The tensor $\mathbf{I}$ is called the *inertia tensor* of the body. It is given by:

$$\mathbf{I} = \sum_\mathbf{r} r^2 \left( \mathbf{I} - \mathbf{r} \mathbf{r} \right) \delta m_\mathbf{r} \quad (1)$$

Show, using Eq.(1), that the angular momentum $\mathbf{L}$ and the kinetic energy $T$ of the body are:

$$\mathbf{L} = \mathbf{I} \cdot \omega \quad (2)$$
$$T = \frac{1}{2} \omega \cdot \mathbf{I} \cdot \omega = \frac{1}{2} \omega_i I_{ij} \omega_j \quad (3)$$

[You will need to expand the vector triple product $\mathbf{f} \times (\mathbf{g} \times \mathbf{h})$ – see Appendix.]

Thus $\mathbf{I}$ specifies the linear dependence of angular momentum $L_i$ on the angular velocity $\omega_j$ (via $L_i = I_{ij} \omega_j$, Eq.(2)) and replaces the single number, the *moment of inertia*, that would be enough for a very symmetrical body, such as a sphere.

(ii) Find the components of $I_{ij}$ in a Cartesian system, that is the usual basis set $\hat{x}, \hat{y}, \hat{z}$.

[Answer: $I_{xx} = \sum_\mathbf{r} (y^2 + z^2) \delta m_\mathbf{r}$, $I_{xy} = -\sum_\mathbf{r} (xy) \delta m_\mathbf{r}$, etc..]

(iii) $I_{ij}$ is a *symmetric tensor*, that is $\mathbf{I} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{I}$ for any vector $\mathbf{v}$, which should be clear from the definition (1). Deduce it also from the matrix of components of $I_{ij}$ found in (ii).

(iv) Take the continuum limit in Eq.(1) to obtain formally, as an integral, the inertia tensor of a body of variable density $\rho(\mathbf{r})$.

(v) Write down the inertia tensor of a body consisting of eight masses $M$ at the corners of
a light cubic frame of side \(a\). Choose the coordinate system most convenient to you. What is the moment of inertia of this cube? Why can we talk of the moment of inertia when we strictly speaking have a tensor?

[In fact, even for an asymmetric body one sometimes speaks of the moment of inertia about some specific axis, \(\hat{n}\), say. This is written as \(\hat{n}.I.\hat{n}\) and gives the ratio of the \(\hat{n}\) component of angular momentum \((L.\hat{n})\) to the \(\hat{n}\) component of angular velocity \((\omega.\hat{n})\). Since in general \(L\) has components which are not parallel to \(\hat{n}\), this does not offer a complete description.]

(vi) **Angular velocity:** A body rotates about an axis through its centre of mass with an angular velocity \(\omega\). The velocity of a point in the body is \(v = \omega \times r\). Show that one can equally well express it as \(v = \underline{\omega} \cdot r\) where \(\underline{\omega}\) is an antisymmetric tensor with components:

\[
\begin{pmatrix}
0 & -\omega_z & \omega_y \\
\omega_z & 0 & -\omega_x \\
-\omega_y & \omega_x & 0
\end{pmatrix}
\]

(vii) Show that any vector \(\mathbf{W}\), constant with respect to axes fixed in the body, changes with time in a space-fixed frame, by an amount \(d\mathbf{W}\) in time \(dt\):

\[
d\mathbf{W} = (\omega dt) \times \mathbf{W}
\]

and hence that if \(\mathbf{W}\) is not body-fixed but has a rate of change \((d\mathbf{W}/dt)_{body}\), the vector \(\mathbf{W}\) in the space-fixed frame obeys

\[
(d\mathbf{W}/dt)_{space} = (d\mathbf{W}/dt)_{body} + \omega \times \mathbf{W}.
\]

B. **Principal Axes and Precession**

We now use the above results to solve simply a difficult problem, the rotational motion of an asymmetric body.
(i) What is the inertia tensor of four point masses $M$ fixed at the corners of a massless square frame of side $a$? What are the principal axes, and principal moments of inertia? Why are two of the axes indeterminate?

This is an example of a nontrivial body (unlike the sphere or cube) where the principal moments of inertia are not all the same. Now $\mathbf{L}$ need not be parallel to $\omega$; for a body whose angular momentum is conserved the result is \textit{free precession} of the angular velocity (see (iii) below).

(ii) It is clear from (1) that in general $\mathbf{I}$ for a body is constant in body-fixed axes, and so for an arbitrary motion the inertia tensor $I_{ij}^{(sp)}$ in space-fixed axes is not constant (unless, trivially, it is proportional to $\delta_{ij}$). Show that in general the motion of a nontrivial body whose angular momentum $\mathbf{L}^{(sp)}$ is conserved (in a space-fixed frame) involves an angular velocity $\omega$ which changes in time. [Hint: consider $L_i^{(sp)} = I_{ij}^{(sp)} \omega_j$.]

(iii) The problem of finding $\omega$ for a body of fixed angular momentum (not parallel to a principal axis) is best solved by adopting a body frame where the inertia tensor is constant. Show from (5) that \textit{in the body frame $\mathbf{L}$} obeys

$$\frac{d\mathbf{L}}{dt} = -\omega \wedge \mathbf{L}$$

(iv) Deduce that

$$I_{ij} \omega_j = -\epsilon_{ijk} \omega_j I_{kl} \omega_l$$

and show that, by taking as the body-fixed frame the principal frame of $I_{ij}$, this can be simplified to

$$I_1 \omega_1 = -\omega_2 \omega_3 (I_3 - I_2)$$

$$I_2 \omega_2 = -\omega_3 \omega_1 (I_1 - I_3)$$

$$I_3 \omega_3 = -\omega_1 \omega_2 (I_2 - I_1)$$

where $I_1$, $I_2$, $I_3$ are the principal moments of inertia.

(v) For the inertia tensor of the square frame found in question (i), $I_1 = I_2$ (such a body
is known as a “symmetric top”). Prove for this case that \( \omega \) precesses in the body frame according to

\[
\omega_3 = \text{const.}; \quad \omega_1 = A \cos(\Omega t); \quad \omega_2 = A \sin(\Omega t)
\]

where \( \Omega = \frac{(I_3-I_1)}{I_1}\omega_3 \), and so confirm that \( \omega \) is not a constant vector unless it happens to point along one of the principal axes.

\[\text{C. Stress and strain tensors.}\]

\[\text{C 1. Consider a small area element } dS \text{ embedded at a point } r \text{ in an isotropic liquid.} \]

If the material is subject to pressure \( P \), there is a normal force on each side of the surface element \( PdS \). These point in opposite directions, so the net force on the surface element is zero – which is as it must be, since the element has no mass and would otherwise be subject to infinite acceleration.

In the case of a solid, there can also be a shear force exerted in equal and opposite directions on the two sides of the element; such a force lies in the plane of the element rather than normal to it. In the general case, we can write

\[
dF = \tau \cdot dS
\]

where \( \tau \) is the stress tensor at the point \( r \). For example the element \( \tau_{xz} \) of \( \tau \) determines the \( x \)-force per unit area on an element of surface with its normal pointing in the \( z \)-direction: \( dF_x = \tau_{xz}dS_z \). For a surface element of general orientation \( dF_x = \tau_{xx}dS_x + \tau_{xy}dS_y + \tau_{xz}dS_z \).

(i) Show that a small box of sides \( dx, dy, dz \) experiences a couple \( (\tau_{xz} - \tau_{zx})dx dy dz \) about the \( y \) axis. The corresponding moment of inertia is of order \( (dx)^2(dz)^2dy \). Explain why this means that \( \tau \) must be symmetric. Show also that, if \( \tau(r) \) is a function of position,
then the net force on a finite piece of material is given by

$$F_i = \int \frac{\partial \tau_{ij}}{\partial r_j} dV \ .$$

where (in a conventional notation) $\frac{\partial}{\partial r_j} \equiv \nabla_j$.

(ii) **Strain:** Suppose a solid material is subjected to a small deformation $\mathbf{r} \to \mathbf{r} + \mathbf{u}(\mathbf{r})$. So long as the deformation depends continuously on position, we may write

$$du_i = \frac{\partial u_i}{\partial r_j} dr_j \ .$$

Let us now decompose the tensor $U_{ij} \equiv \frac{\partial u_i}{\partial r_j}$ into its symmetric and anti-symmetric parts:

$$U_{ij} = e_{ij} + R_{ij}$$

$$e_{ij} = \frac{1}{2} (U_{ij} + U_{ji})$$

$$R_{ij} = \frac{1}{2} (U_{ij} - U_{ji})$$

Show that $e$ is symmetric and $R$ antisymmetric. Show also that $R$ corresponds to a pure rotation (i.e. without distortion) of the bit of matter near $\mathbf{r}$. [Hint: compare Eq.(4) in section A.] What type of deformation does $e$ represent? Why may we assume that the stress tensor $\tau$ depends on the strain tensor $e$, but has no dependence on $R$?

(iii) **Elasticity:** For small deformations, we expect the stress and strain to be linearly related, but since both are described by second-rank tensors, the relation between them generally involves a tensor of the fourth rank:

$$\tau_{ij} = E_{ijkl} e_{kl} \tag{10}$$

which is the generalization of Hooke's law to an arbitrary three-dimensional material. Since $\tau$ is symmetric, $E_{ijkl}$ must be symmetric under exchange of $i$ and $j$; various other symmetries can be found from considering the form of the stored elastic energy, but even when these are taken into account, $E_{ijkl}$ has 21 independent components.
Isotropic elastic solids are much simpler however, and have only two independent constants. These are the Young’s modulus \(Y\) and Poisson’s ratio \(\sigma\). The definition of these is that in the principal axes of \(\tau_{ij}\), the following relation applies

\[ Ye_{xx} = \tau_{xx} - \sigma(\tau_{yy} + \tau_{zz}) \]  

with similar relations for \(e_{yy}, e_{zz}\); all off-diagonal elements of \(e\) are zero in this coordinate system. By rearranging Eq.(11) obtain the result

\[ Ye_{xx} = (1 + \sigma)\tau_{xx} - \sigma \tau_{ii} \]

(where \(x\) indices are not summed over but \(\tau_{ii}\) is summed) with similar relations for \(yy\) and \(zz\) components. [Note that \(\tau_{ii} = \text{Trace}(\tau)\) under the summation convention.] Show that, by adding Eq.(11) to its analogues,

\[ \tau_{ii} = \frac{Ye_{ii}}{1 - 2\sigma} \]

Combine the above two results to give expressions for \(\tau_{xx}\) etc in terms of the elements of \(e\), and hence derive the following result in the principal frame:

\[ \tau_{ij} = \frac{Y}{1 + \sigma} \left( e_{ij} + \frac{\sigma}{1 - 2\sigma} \delta_{ij} \text{Tr}(e) \right). \]

[Bear in mind that \(e_{ij} = 0\) for \(i \neq j\), in this frame.] Hence, derive the general (frame independent) form for the elasticity tensor in an isotropic solid:

\[ E_{ijkl} = \frac{Y}{1 + \sigma} \left[ \delta_{ik} \delta_{jl} + \frac{\sigma}{2} \delta_{ij} \delta_{kl} + \left( \frac{\sigma}{1 - 2\sigma} \right) \delta_{ij} \delta_{kl} \right]. \]

C 2. (i) Apply \(E_{ijkl}\) to suitable deformations to recover expressions for the bulk modulus and shear modulus of an isotropic solid in terms of \(Y\) and \(\sigma\).

(ii) Show that the displacement \(u\) inside a homogeneous isotropic solid, subject to external forces only at its surface, obeys

\[ \text{grad div}u + (1 - 2\sigma) \nabla^2 u = 0. \]
D. Transformation Laws.

D 1. Here we consider transformations of vectors and tensors between two sets of base vectors, each of which is orthonormal. We call our two sets of base vectors \( \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \) and \( \mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \), so that we can use suffix notation, e.g., the fact that the \( \mathbf{e} \) basis is orthonormal can be written

\[
\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \tag{12}
\]

The unit vectors \( \mathbf{e}'_i \) are linear combinations of the \( \mathbf{e}_i \):

\[
\mathbf{e}'_i = T_{ij} \mathbf{e}_j . \tag{13}
\]

Suppose that \( \mathbf{x} = x_i \mathbf{e}_i = x'_i \mathbf{e}'_i \) is any vector, and \( \mathbf{y} = A \mathbf{x} \). Prove that

(i) \( \mathbf{e}'_i \cdot \mathbf{e}_k = T_{ik} \)

(ii) The inverse transformation \( \mathbf{e}_i = T^{-1}_{ik} \mathbf{e}'_k \) has \( T^{-1}_{ik} = T_{ki} \)

(iii) \( x'_i = T_{ik} x_k \)

(iv) \( x_i = T_{ki} x'_k \)

(v) \( y'_i = T_{ij} T_{lk} A_{jk} x'_l \)

This means you have found the components of the tensor \( A \) in the rotated coordinate system:

\[
A'_{il} = T_{ij} T_{lk} A_{jk} \]

Indeed, this property under orthogonal coordinate transformations \( T \) is, mathematically, often taken as the defining property of a tensor \( A \). For tensors of higher rank than 2, one factor of the transformation matrix \( T \) is needed for each index. Transformations with \( \det(T) = 1 \) are usually called proper rotations and those with \( \det(T) = -1 \), improper rotations. Show that these are the only possible values of the determinant for transformations between orthonormal bases as defined above.
Appendix: The antisymmetric three-tensor $\epsilon_{ijk}$.

[Reprise of Sheet V section D.]

A tensor with two indices such as $\delta_{ij}$ is called a tensor of second rank, and its linear operation on a vector gives another vector. A vector is technically a tensor of rank one; a scalar of rank zero. Tensors of higher rank can also be defined. For example a tensor of rank four, $Q_{ijkl}$ acts on a rank two tensor to give another rank two tensor, $B_{ij} = Q_{ijkl}A_{kl}$ (with summation over repeated indices). For tensors higher than rank two, dyadic notation becomes extremely cumbersome and suffix notation is essential.

1. An important and useful tensor of rank three is the fully antisymmetric three tensor denoted $\epsilon_{ijk}$. This operates on either a second rank tensor or two vectors to give a vector. For $i \neq j \neq k$ it is defined as $\epsilon_{ijk} = 1$ if $ijk$ is a cyclic permutation of $x, y, z$ (e.g., $yzx$) and $= -1$ if the order is anticyclic (e.g., $xzy$). If any two of $ijk$ are the same then $\epsilon_{ijk} = 0$ (e.g., $\epsilon_{xxy} = 0$). The tensor $\epsilon_{ijk}$ is very useful for writing down a vector product in suffix notation. Show that

$$\mathbf{A} \wedge \mathbf{B} = \epsilon_{ijk} A_j B_k$$

(14)

The following relation

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

(15)

is tedious to prove (using components – you should come across this in maths lectures) but is worth remembering. It saves a great deal of time when manipulating vector identities such as

$$\mathbf{A} \wedge (\mathbf{B} \wedge \mathbf{C}) = \epsilon_{ijk} A_j \epsilon_{klm} B_k C_m$$

$$= A_j B_k C_m (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl})$$

$$= B_i A_j C_j - C_i A_j B_j = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C}$$
which is a familiar result normally derived by more tedious means.

The result Eq.(15) is especially helpful when dealing with vector identities involving the $\nabla$ operator. Because $\epsilon_{ijk}$ is a constant tensor, it commutes with $\nabla$, and expressions can be re-ordered accordingly. For example

$$\nabla \cdot (\vec{E} \wedge \vec{H}) \equiv \nabla \epsilon_{ijk} E_j H_k$$

(16)

$$= \epsilon_{ijk} \nabla E_j H_k = \epsilon_{ijk}[E_j \nabla_i H_k + H_k \nabla_i E_j]$$

(17)

$$= -E_j \epsilon_{ijk} \nabla H_k + H_k \epsilon_{kij} \nabla E_j$$

(18)

$$= -\vec{E} \cdot \nabla \wedge \vec{H} + \vec{H} \cdot \nabla \wedge \vec{E}$$

(19)

which is a result used in connection with Poynting’s vector in electromagnetism.

(i) Make sure you understand each step of this derivation.

(ii) Show, using similar methods, that

$$\nabla \wedge (\nabla \wedge \vec{A}) = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A}$$

$$\nabla \wedge (\psi \vec{A}) = \psi \nabla \wedge \vec{A} - \vec{A} \wedge \nabla \psi$$

These relations are useful in electromagnetism and other topics in the IB and parts II and III courses.