EXAMPLES CLASS IN MATHEMATICAL PHYSICS

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 ϵ_{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 $\underline{P}$ induced in a medium by applying a field $\underline{E}$

$$\underline{P} = \epsilon_0 \chi \underline{E} \tag{1}$$

Here $\underline{P}$ and $\underline{E}$ 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 $\underline{P}$ and $\underline{E}$ 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, $p_{||} = \beta_{||} E_{||}$ being the induced dipole for a field $E_{||}$ applied along the molecule. The response to perpendicular fields is $p_\perp = \beta_\perp E_\perp$ where $\beta_\perp \ll \beta_{||}$. Let the cylindrical molecule point in the $(1, 1, 0)$ direction and apply a field in the x direction: $\underline{E} = E_x \underline{\hat{x}}$

(i) Prove:

$$p_x = (\beta_{||} + \beta_\perp) E_x / 2 \tag{2}$$

$$p_y = (\beta_{||} - \beta_\perp) E_x / 2 \tag{3}$$

$$p_z = 0 \tag{4}$$

Hence one sees that $\underline{p}$ and $\underline{E}$ are not parallel and in general we must write

$$\underline{p} = \underline{\underline{\alpha}} \cdot \underline{E} \tag{5}$$

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

$$p_x = \alpha_{xx}E_x + \alpha_{xy}E_y + \alpha_{xz}E_z \tag{6}$$

$$p_y = \alpha_{yx}E_x + \alpha_{yy}E_y + \alpha_{yz}E_z \tag{7}$$

$$p_z = \alpha_{zx}E_x + \alpha_{zy}E_y + \alpha_{zz}E_z \tag{8}$$

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

Remark: If many molecules are present with the same orientation, $\underline{P}$ is $\rho\underline{p}$ where there are ρ 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 $\underline{\underline{\chi}}$ (to replace χ in (1)) would (neglecting internal field corrections) be $\rho\underline{\underline{\alpha}}$.

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 $\underline{\underline{\alpha}}$ can be denoted α_{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* δ_{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

$$Q_{ij} \equiv a_i \, b_j \tag{9}$$

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

(i) Show using suffix notation that if $\underline{p}$ is a vector, $\underline{\underline{Q}} \cdot \underline{p}$ is also a vector; find its magnitude and direction. [Note that $\underline{\underline{Q}} \cdot \underline{p} \equiv Q_{ij} p_j$.] We see that Q_{ij} is a linear operator that transforms one vector $\underline{p}$ 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

$$\mathbf{Q} = \mathbf{ab} \tag{10}$$

or $\underline{\underline{Q}} = \underline{a}\,\underline{b}$, 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 $\underline{\underline{\alpha}}$ in Q. A1 in the following form:

$$\underline{\underline{\alpha}} = \alpha_{xx} \, \underline{\hat{x}} \, \underline{\hat{x}} + \alpha_{xy} \, \underline{\hat{x}} \, \underline{\hat{y}} + \alpha_{xz} \, \underline{\hat{x}} \, \underline{\hat{z}} + 6 \text{ other terms} \tag{11}$$

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

$$\mathbf{p} = \underline{\underline{\alpha}} \cdot \mathbf{E} \equiv \underline{\underline{\alpha}} \cdot \hat{\mathbf{x}} \, E_x = (\alpha_{xx}, \alpha_{yx}, \alpha_{zx}) \, E_x \ .$$

(iii) **Projection operators**: In dyadic notation, the identity tensor δ_{ij} is often written as the "unit operator" $\underline{\underline{1}}$. Consider the action on a vector $\mathbf{v}$ of the operators $\hat{\mathbf{u}}\hat{\mathbf{u}}$ and $\underline{\underline{1}} - \hat{\mathbf{u}}\hat{\mathbf{u}}$, where $\hat{\mathbf{u}}$ is a unit vector. Show that these perform the job of resolving the vector $\mathbf{v}$ into vectors $\mathbf{v}_{||}$ and $\mathbf{v}_{\perp}$, respectively parallel and perpendicular to $\hat{\mathbf{u}}$. 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

$$\underline{\underline{\alpha}} = \beta_{||}\hat{\mathbf{u}}\hat{\mathbf{u}} + \beta_{\perp}(\underline{\underline{1}} - \hat{\mathbf{u}}\hat{\mathbf{u}})$$

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

C. 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 1×10^{-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 $\hat{\mathbf{u}}$,$\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$ for which there is no torque on the object; what relation do $\mathbf{E}$ and $\mathbf{p}$ then have? Show that $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$ are the eigenvectors of $\underline{\underline{\alpha}}$.

(iv) Show that, if we use as a coordinate system the unit vectors $\hat{\mathbf{u}}$,$\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$ 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, λ_1, λ_2 and λ_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 $\hat{\mathbf{u}}$,$\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$ in terms of $\underline{\hat{x}}$, $\underline{\hat{y}}$, $\underline{\hat{z}}$?

(v) In this example, $\underline{\underline{\alpha}}$ 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 $\hat{\mathbf{u}}$,$\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$ 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 $\underline{\underline{\alpha}}$ into principal directions $\hat{\mathbf{u}}$,$\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$ can be accomplished using projection operators:

$$\underline{\underline{\alpha}} = \lambda_1 \hat{\mathbf{u}}\hat{\mathbf{u}} + \lambda_2 \hat{\mathbf{v}}\hat{\mathbf{v}} + \lambda_3 \hat{\mathbf{w}}\hat{\mathbf{w}} \tag{12}$$

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 $\underline{\underline{\alpha}}$ on a vector $\underline{b}$. To find $\underline{\underline{\alpha}} \cdot \underline{b}$, the recipe is: (1) resolve $\underline{b}$ 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 $\underline{\underline{\alpha}}$ on $\underline{b}$ 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

$$F_i = \frac{\partial E_i}{\partial x_j} p_j \tag{13}$$

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 $\mathbf{F}$ =grad($\mathbf{p}.\mathbf{E}$), 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 $\mathbf{E}$ 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 $F = \alpha\ \mathrm{grad}(E^2/2)$?

D. The antisymmetric three-tensor ϵ_{ijk}.

A tensor with two indices such as δ_{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 ϵ_{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 ϵ_{ijk} is very useful for writing down a vector product in suffix notation. Show that

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

The following relation

$$\epsilon_{ijk}\, \epsilon_{klm} = \delta_{il}\, \delta_{jm} - \delta_{im}\, \delta_{jl} \tag{15}$$

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

$$\underline{A} \wedge (\underline{B} \wedge \underline{C}) = \epsilon_{ijk}\, A_j\, \epsilon_{klm}\, B_l\, C_m$$

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

$$= B_i\, A_j\, C_j - C_i\, A_j\, B_j = (\underline{A} \cdot \underline{C})\underline{B} - (\underline{A} \cdot \underline{B})\underline{C}$$

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 ∇ operator. Because ϵ_{ijk} is a constant tensor, it commutes with ∇, and expressions can be re-ordered accordingly. For example

$$\nabla \cdot (\underline{E} \wedge \underline{H}) \equiv \nabla_i(\epsilon_{ijk}\, E_j\, H_k) \tag{16}$$

$$= \epsilon_{ijk}\nabla_i(E_j\, H_k) = \epsilon_{ijk}[E_j \nabla_i H_k + H_k \nabla_i E_j] \tag{17}$$

$$= -E_j\, \epsilon_{jik}\nabla_i H_k + H_k\, \epsilon_{kij}\nabla_i E_j \tag{18}$$

$$= -\underline{E} \cdot \nabla \wedge \underline{H} + \underline{H} \cdot \nabla \wedge \underline{E} \tag{19}$$

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

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

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

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