Tensors
(Symon Chapter Ten)

Physics A301*

Spring 2004

Contents

1 Angular Momentum and Angular Velocity

So far we know about scalars (single number) and vectors (which can be written as triples of components). In rigid body motion, we will want to deal with more complicated objects. The relationship between the angular momentum $\vec{L}$ and angular velocity $\vec{\omega}$ will be described by a tensor $\overleftrightarrow{I}$.

First recall momentum and velocity. Imagine N particles all moving with the same velocity $\vec{v}$. The total momentum is

$$\vec{P} = \sum_{k=1}^{N} \vec{p}_k = \sum_{k=1}^{N} m_k \vec{v} = \underbrace{\left(\sum_{k=1}^{N} m_k \right)}_{M} \vec{v} = M\vec{v} \tag{1.1}$$

The momentum vector equals a scalar mass times the velocity vector. So

- $\vec{P} \parallel \vec{v}$
- $\vec{P}$ is linear in $\vec{v}$, i.e., if $\vec{v} \longrightarrow c\vec{v}$, then $\vec{P} \longrightarrow c\vec{P}$.

Now, consider N particles all rotating with angular velocity $\vec{\omega}$. From chapter 7, we recall this means each one has a velocity

$$\vec{v}_k = \dot{\vec{r}}_k = \vec{\omega} \times \vec{r}_k \tag{1.2}$$

The total angular momentum is

$$\begin{aligned} \vec{L} &= \sum_{k=1}^{N} \vec{r}_k \times \vec{p}_k = \sum_{k=1}^{N} \vec{r}_k \times (m_k \dot{\vec{r}}_k) = \sum_{k=1}^{N} m_k \overbrace{\vec{r}_k \times (\vec{\omega} \times \vec{r}_k)}^{=(\vec{r}_k \cdot \vec{r}_k)\vec{\omega} - \vec{r}_k(\vec{r}_k \cdot \vec{\omega})} \\ &= \sum_{k=1}^{N} m_k \big(r_k^2 \vec{\omega} - \vec{r}_k \underbrace{(\vec{r}_k \cdot \vec{\omega})}_{\text{scalar}} \big) \end{aligned} \tag{1.3}$$

Note that $\vec{L}$ is still linear in $\vec{\omega}$ (if $\vec{\omega} \longrightarrow c\vec{\omega}$, then $\vec{L} \longrightarrow c\vec{L}$) but the two are not in general parallel. We'd like to pull $\vec{\omega}$ out of the sum like we did with $\vec{v}$ but we don't yet have a notation for it.

1.1 Matrix Approach

We *can* write the components in a matrix:

$$\begin{aligned} \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} &= \sum_{k=1}^{N} m_k \left[r_k^2 \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} - \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} (x_k\omega_x + y_k\omega_y + z_k\omega_z) \right] \\ &= \sum_{k=1}^{N} m_k \left[r_k^2 \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} - \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} \begin{pmatrix} x_k & y_k & z_k \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \right] \end{aligned} \tag{1.4}$$

Well, that's interesting. If we define column vectors (3×1 matrices)[1] $\mathbf{L}$, $\mathbf{r}_k$ and $\boldsymbol{\omega}$, (1.4) can be written

$$\mathbf{L} = \sum_{k=1}^{N} \left[r_k^2 \boldsymbol{\omega} - \mathbf{r}_k ({\mathbf{r}_k}^{\mathrm{T}} \boldsymbol{\omega}) \right] \tag{1.5}$$

where ${\mathbf{r}_k}^{\mathrm{T}}$ is the 1×3 matrix which is the transpose of $\mathbf{r}_k$. But, now, matrix multiplication is associative, so

$$\mathbf{r}_k ({\mathbf{r}_k}^{\mathrm{T}} \boldsymbol{\omega}) = (\mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}}) \boldsymbol{\omega} \tag{1.6}$$

where the quantity in parentheses on the right-hand side is the 3×3 matrix

$$\mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}} = \begin{pmatrix} x_k x_k & x_k y_k & x_k z_k \\ y_k x_k & y_k y_k & y_k z_k \\ z_k x_k & z_k y_k & z_k z_k \end{pmatrix} \tag{1.7}$$

So ...

$$\mathbf{L} = \sum_{k=1}^{N} m_k \left[r_k^2 \boldsymbol{\omega} - (\mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}}) \boldsymbol{\omega} \right] = \left\{ \sum_{k=1}^{N} m_k \left[r_k^2 \mathbf{1} - \mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}} \right] \right\} \boldsymbol{\omega} = \mathbf{I} \boldsymbol{\omega} \tag{1.8}$$

where

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.9}$$

is the identity matrix and we've defined the 3×3 matrix

$$\begin{aligned} \mathbf{I} &= \sum_{k=1}^{N} m_k \left[r_k^2 \mathbf{1} - \mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}} \right] \\ &= \sum_{k=1}^{N} m_k \begin{pmatrix} x_k^2 + y_k^2 + z_k^2 & 0 & 0 \\ 0 & x_k^2 + y_k^2 + z_k^2 & 0 \\ 0 & 0 & x_k^2 + y_k^2 + z_k^2 \end{pmatrix} - \begin{pmatrix} x_k x_k & x_k y_k & x_k z_k \\ y_k x_k & y_k y_k & y_k z_k \\ z_k x_k & z_k y_k & z_k z_k \end{pmatrix} \end{aligned} \tag{1.10}$$

2 Tensors

Of course, we don't do physics in terms of matrices, but rather in terms of *vectors* to stress the geometrical significance.[2]

The operation

$${\mathbf{r}_k}^{\mathrm{T}} \boldsymbol{\omega} = x_k \omega_x + y_k \omega_y + z_k \omega_z = \vec{r}_k \cdot \vec{\omega} \tag{2.1}$$

we understand as a dot (scalar) product. But how about $\mathbf{r}_k {\mathbf{r}_k}^{\mathrm{T}}$ which would produce not a number but a 3×3 matrix?

[1] We use boldfaced letters to describe matrices but *not* vectors, as defined in the notational handout.

[2] A matrix $\mathbf{v}$ is just a particular set of three numbers which may represent the components of a vector in one basis. A vector $\vec{v}$ is a geometric object which has meaning independent of any particular choice of basis.

2.1 The Tensor (dyad) Product

We define a new notation. $\vec{r} \otimes \vec{r}$ is the **Tensor Product** of a vector with itself.[3]

In general, the object $\vec{A} \otimes \vec{B}$ is a "dyad", which is a kind of *tensor*. (In general a tensor is a dyad or a sum of dyads.) It is defined by its dot product with a vector $\vec{C}$, i.e.,

$$(\vec{A} \otimes \vec{B}) \cdot \vec{C} = \vec{A}(\vec{B} \cdot \vec{C}) \tag{2.2}$$

We define another kind of tensor, the identity tensor $\overleftrightarrow{1}$ by its dot product $\overleftrightarrow{1} \cdot \vec{C} = \vec{C}$.

With this expansion of notation we can write

$$\vec{L} = \overleftrightarrow{I} \cdot \vec{\omega} \tag{2.3}$$

where

$$\overleftrightarrow{I} = \sum_{k=1}^{N} m_k (r_k^2 \overleftrightarrow{1} - \vec{r}_k \otimes \vec{r}_k) \tag{2.4}$$

is the *inertia tensor*, a property of the mass distribution.

2.2 Tensor Operation

Given a tensor $\overleftrightarrow{T}$ we can define the following operations:

- Multiplication by a scalar $a\overleftrightarrow{T}$
- Addition of two tensors $\overleftrightarrow{S} + \overleftrightarrow{T}$
- Dot product with a vector $\overleftrightarrow{T} \cdot \vec{A}$ or $\vec{A} \cdot \overleftrightarrow{T}$ (note these are in general *not* the same.

As a warning, Symon's equation (10.19) is trying to say that $\overleftrightarrow{T} \cdot \vec{C} \neq \vec{C} \cdot \overleftrightarrow{T}$ in general, but if you only read the equation and not the text around it, it looks like the opposite. As an explicit example, consider the case where $\overleftrightarrow{T} = \hat{x} \cdot \hat{y}$ and $\vec{C} = \hat{x}$. Then

$$\overleftrightarrow{T} \cdot \vec{C} = (\hat{x} \cdot \hat{y}) \cdot \hat{x} = \hat{x}\cancelto{0}{(\hat{y} \cdot \hat{x})} = \vec{0} \tag{2.5}$$

but

$$\overleftrightarrow{T} \cdot \vec{C} = \hat{x} \cdot (\hat{x} \cdot \hat{y}) = (\hat{x} \cdot \hat{x})\hat{y} = \hat{y} \tag{2.6}$$

2.3 Components of a Tensor

Just as any vector can be understood in terms of its components in an orthonormal basis:

$$\vec{C} = C_x \hat{x} + C_y \hat{y} + C_z \hat{z} \tag{2.7}$$

any tensor also has components.

[3] In Symon's notation, the two vectors are just written side-by-side, so he calls this **rr**. This is potentially confusing notation, since it's the same as ordinary scalar multiplication, so we will *always* include the "$\otimes$" and you're expected to do likewise on homeworks and exams.

Start with a dyad:

$$\begin{aligned}\vec{A}\otimes\vec{B} &= (A_x\,\hat{x} + A_y\,\hat{y} + A_z\,\hat{z})\otimes(B_x\,\hat{x} + B_y\,\hat{y} + B_z\,\hat{z}) \\ &= \quad A_xB_x\;\hat{x}\otimes\hat{x} + A_xB_y\;\hat{x}\otimes\hat{y} + A_xB_z\;\hat{x}\otimes\hat{z} \\ &\quad + A_yB_x\;\hat{y}\otimes\hat{x} + A_yB_y\;\hat{y}\otimes\hat{y} + A_yB_z\;\hat{y}\otimes\hat{z} \\ &\quad + A_zB_x\;\hat{z}\otimes\hat{x} + A_zB_y\;\hat{z}\otimes\hat{y} + A_zB_z\;\hat{z}\otimes\hat{z}\end{aligned} \tag{2.8}$$

In general, any tensor can be written

$$\begin{aligned}\overleftrightarrow{T} = &\quad T_{xx}\;\hat{x}\otimes\hat{x} + T_{xy}\;\hat{x}\otimes\hat{y} + T_{xz}\;\hat{x}\otimes\hat{z} \\ &+ T_{yx}\;\hat{y}\otimes\hat{x} + T_{yy}\hat{y}\otimes\hat{y} + T_{yz}\;\hat{y}\otimes\hat{z} \\ &+ T_{zx}\;\hat{z}\otimes\hat{x} + T_{zy}\;\hat{z}\otimes\hat{y} + T_{zz}\;\hat{z}\otimes\hat{z}\end{aligned} \tag{2.9}$$

2.3.1 Tensor Components and the Matrix Representation

Now we can make the connection to the matrix $\mathbf{T}$ in the analogy; if

$$\begin{aligned}\vec{X} = \overleftrightarrow{T}\cdot\vec{C} = \big[&(T_{xx}\,\hat{x} + T_{yx}\,\hat{y} + T_{zx}\,\hat{z})\otimes\hat{x} + (T_{xy}\,\hat{x} + T_{yy}\,\hat{y} + T_{zy}\,\hat{z})\otimes\hat{y} \\ &+ (T_{xz}\,\hat{x} + T_{yz}\,\hat{y} + T_{zz}\,\hat{z})\otimes\hat{z}\big]\cdot(C_x\,\hat{x} + C_y\,\hat{y} + C_z\,\hat{z}) \\ = (T_{xx}C_x + T_{xy}C_y + T_{xz}C_z)\hat{x} &+ (T_{yx}C_x + T_{yy}C_y + T_{yz}C_z)\hat{y} + (T_{zx}C_x + T_{zy}C_y + T_{zz}C_z)\hat{z}\end{aligned} \tag{2.10}$$

then

$$\mathbf{X} = \begin{pmatrix} T_{xx}C_x + T_{xy}C_y + T_{xz}C_z \\ T_{yx}C_x + T_{yy}C_y + T_{yz}C_z \\ T_{zx}C_x + T_{zy}C_y + T_{zz}C_z \end{pmatrix} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix}\begin{pmatrix} C_x \\ C_y \\ C_z \end{pmatrix} = \mathbf{TC} \tag{2.11}$$

To make the notation a little more compact, define

$$x_1 = x \tag{2.12a}$$

$$x_2 = y \tag{2.12b}$$

$$x_3 = z \tag{2.12c}$$

and

$$\hat{e}_1 = \hat{x} \tag{2.13a}$$

$$\hat{e}_2 = \hat{y} \tag{2.13b}$$

$$\hat{e}_3 = \hat{z} \tag{2.13c}$$

so that

$$\vec{C} = \sum_{i=1}^{3} C_i\hat{e}_i \tag{2.14a}$$

$$\overleftrightarrow{T} = \sum_{i=1}^{3}\sum_{j=1}^{3} T_{ij}\hat{e}_i\otimes\hat{e}_j \tag{2.14b}$$

$$\overleftrightarrow{T}\cdot\vec{C} = \sum_{i=1}^{3}\sum_{j=1}^{3}(T_{ij}C_j)\hat{e}_i \tag{2.14c}$$

Note the repetition of the index j on $T_{ij}C_j$; this is the standard pattern for matrix multiplication.

2.4 Inertia Tensor of a Solid Body

Recall our definition of the inertia tensor

$$\overleftrightarrow{I} = \sum_{k=1}^{N} m_k (r_k^2 \overleftrightarrow{1} - \vec{r}_k \otimes \vec{r}_k) \tag{2.15}$$

This is for a collection of point masses. For a continuous mass distribution, we make the usual substitution

$$\sum_{k=1}^{N} m_k \longrightarrow \iiint \rho(\vec{r})\, d^3V \tag{2.16}$$

so

$$\overleftrightarrow{I} = \iiint \rho(\vec{r}) \left[r^2 \overleftrightarrow{1} - \vec{r} \otimes \vec{r} \right] d^3V \tag{2.17}$$

We can write this in terms of components as

$$I_{ij} = \iiint \rho(\vec{r}) \left[r^2 \delta_{ij} - r_i r_j \right] d^3V \tag{2.18}$$

Where we have defined the incredibly useful notation for the components of the identity tensor:

$$\overleftrightarrow{1} = \sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij}\ \hat{e}_i \otimes \hat{e}_j \tag{2.19}$$

or of the identity matrix:

$$\mathbf{1} = \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.20}$$

so

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{2.21}$$

This is called the *Kronecker delta* and is incredibly useful. Learn it!

3 More Properties of Tensors

3.1 Transpose

A vector $\vec{C}$ can correspond to a column vector $\mathbf{C}$ or a row vector $\mathbf{C}^{\mathrm{T}}$, e.g., in

$$\vec{A} \cdot \vec{B} = \mathbf{A}^{\mathrm{T}} \mathbf{B} \tag{3.1}$$

and

$$\vec{A} \otimes \vec{B} \longleftrightarrow \mathbf{A}\mathbf{B}^{\mathrm{T}} \tag{3.2}$$

Given a tensor $\overleftrightarrow{T}$ which has a corresponding 3×3 matrix $\mathbf{T}$, we can take the transpose $\overleftrightarrow{\mathbf{T}}$ which is a *different* 3×3 matrix and ask what tensor corresponds to that.

In terms of components

$$T_{ij}{}^{\mathrm{T}} = T_{ji} \tag{3.3}$$

so

$$\overleftrightarrow{T}^{\mathrm{T}} = \sum_{i=1}^{3}\sum_{j=1}^{3} T_{ij}{}^{\mathrm{T}} \hat{e}_i \otimes \hat{e}_j = \sum_{i=1}^{3}\sum_{j=1}^{3} T_{ji} \hat{e}_i \otimes \hat{e}_j = \sum_{i=1}^{3}\sum_{j=1}^{3} T_{ij} \underbrace{\hat{e}_j \otimes \hat{e}_i}_{(\hat{e}_j \otimes \hat{e}_i)^{\mathrm{T}}} \tag{3.4}$$

where we've renamed i to j and vice-versa in the last step.

Another way to get this is to start from

$$(\vec{A} \otimes \vec{B})^{\mathrm{T}} = \vec{B} \otimes \vec{A} \tag{3.5}$$

A symmetric tensor obeys $\overleftrightarrow{T}^{\mathrm{T}} = \overleftrightarrow{T}$, i.e., $T_{ij} = T_{ji}$.

An antisymmetric tensor obeys $\overleftrightarrow{T}^{\mathrm{T}} = -\overleftrightarrow{T}$, i.e., $T_{ij} = -T_{ji}$.

Examples of symmetric tensors include the identity tensor $\overleftrightarrow{1}$, the dyad product of a vector with itself $\vec{A} \otimes \vec{A}$, and the inertia tensor $\overleftrightarrow{I}$.

As an example of an antisymmetric tensor, start with any vector $\vec{A}$ and define a tensor $\overleftrightarrow{{}^*A}$ which corresponds to the matrix

$${}^*\mathbf{A} = \begin{pmatrix} 0 & A_z & -A_y \\ -A_z & 0 & A_x \\ A_y & -A_x & 0 \end{pmatrix} \tag{3.6}$$

Note that

$${}^*\mathbf{A}\,\mathbf{B} = \begin{pmatrix} 0 & A_z & -A_y \\ -A_z & 0 & A_x \\ A_y & -A_x & 0 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = \begin{pmatrix} A_z B_y - A_y B_z \\ -A_z B_x + A_x B_z \\ A_y B_x - A_x B_y \end{pmatrix} \tag{3.7}$$

So

$$\overleftrightarrow{{}^*A} \cdot \vec{B} = -\vec{A} \times \vec{B} \tag{3.8}$$

If we count the number of independent components, a symmetric tensor has six and an antisymmetric tensor had three, the same number as a vector.

$$\text{Symmetic:} \quad \begin{pmatrix} 1 & 4 & 5 \\ \cdot & 2 & 6 \\ \cdot & \cdot & 3 \end{pmatrix} \qquad \text{Antisymmetic:} \quad \begin{pmatrix} \times & 1 & 2 \\ \cdot & \times & 3 \\ \cdot & \cdot & \times \end{pmatrix} \tag{3.9}$$

So in fact *any* antisymmetric tensor can be written as the dual of a vector.

3.2 Change of Basis

The same vector can be written in two different bases:

$$\vec{v} = v_x\,\hat{x} + v_y\,\hat{y} + v_z\,\hat{z} = v_{x'}\,\hat{x}' + v_{y'}\,\hat{y}' + v_{z'}\,\hat{z}' = \sum_{i=1}^{3} v_i \hat{e}_i = \sum_{i=1}^{3} v_{i'} \hat{e}'_i \tag{3.10}$$

The same can be done with a tensor:

$$\overleftrightarrow{T} = \sum_{i=1}^{3}\sum_{j=1}^{3} \overbrace{T_{ij}}^{\substack{\text{elements of}\\ \text{matrix } \mathbf{T}}} \;\hat{e}_i \otimes \hat{e}_j = \sum_{i=1}^{3}\sum_{j=1}^{3} \overbrace{T_{i'j'}}^{\substack{\text{elements of}\\ \text{matrix } \mathbf{T}'}} \;\hat{e}'_i \otimes \hat{e}'_j \tag{3.11}$$

Note

$$T_{k'\ell'} = \hat{e}'_k \cdot \overleftrightarrow{T} \cdot \hat{e}'_\ell = \sum_{i=1}^{3}\sum_{j=1}^{3} \underbrace{(\hat{e}'_k \cdot \hat{e}_i)}_{A_{ki}} T_{ij} \underbrace{(\hat{e}_j \cdot \hat{e}'_\ell)}_{A_{\ell j} = A_{j\ell}{}^{\mathrm{T}}} \tag{3.12}$$

This is equivalent to the matrix expression

$$\mathbf{T}' = \mathbf{A}\mathbf{T}\mathbf{A}^{\mathrm{T}} \tag{3.13}$$

Note

$$\overleftrightarrow{1} = \sum_{i=1}^{3} \hat{e}_i \otimes \hat{e}_i = \sum_{i=1}^{3} \hat{e}'_i \otimes \hat{e}'_i \tag{3.14}$$

is the same in any orthonormal basis because the transformation between orthonormal bases obeys

$$\mathbf{A}\,\mathbf{A}^{\mathrm{T}} = \mathbf{1} \tag{3.15}$$

i.e.,

$$A_{ik}A_{jk} = \delta_{ij} \tag{3.16}$$

We call $\mathbf{A}$ an “orthogonal” matrix.

3.2.1 Diagonalization of a Symmetric Tensor

Claim: Given any symmetric tensor $\overleftrightarrow{T}$, there is an orthonormal basis $\{\hat{u}_i | i = 1 \ldots 3\}$ such that we can write

$$\overleftrightarrow{T} = \sum_{i=1}^{3} T_i\,\hat{u}_i \otimes \hat{u}_i = T_1\,\hat{u}_1 \otimes \hat{u}_1 + T_2\,\hat{u}_2 \otimes \hat{u}_2 + T_3\,\hat{u}_3 \otimes \hat{u}_3 \tag{3.17}$$

This is because the corresponding symmetric matrix can be diagonalized.

Recall the eigenvalue problem: for which column vectors $\mathbf{v}$ and numbers λ does $\mathbf{T}\mathbf{v} = \lambda\mathbf{v}$? First, write it as

$$(\mathbf{T} - \lambda\mathbf{1})\mathbf{v} = \mathbf{0} \tag{3.18}$$

If $\mathbf{T} - \lambda\mathbf{1}$ is invertible, the solution is

$$\mathbf{v} = (\mathbf{T} - \lambda\mathbf{1})^{-1}\mathbf{0} = \mathbf{0} \tag{3.19}$$

so there is only a non-trivial solution if $\det(\mathbf{T} - \lambda\mathbf{1}) = 0$:

$$\begin{vmatrix} T_{11}-\lambda & T_{12} & T_{13} \\ T_{21} & T_{22}-\lambda & T_{23} \\ T_{31} & T_{32} & T_{33}-\lambda \end{vmatrix} = 0 \tag{3.20}$$

Since the left-hand side is a cubic polynomial in λ, there are (up to) three eigenvalues.

- If $\mathbf{T}$ is real and symmetric, its eigenvalues are all real numbers.
- If $\mathbf{T}$ is real and symmetric, eigenvectors corresponding to different eigenvalues are perpendicular.
- If multiple eigenvectors have the same eigenvalue, any linear combination of them is also an eigenvector with the same eigenvalue:

$$\mathbf{T}(a\mathbf{v}_1 + b\mathbf{v}_2) = a\lambda\mathbf{v}_1 + b\lambda\mathbf{v}_2 = \lambda(a\mathbf{v}_1 + b\mathbf{v}_2) \tag{3.21}$$

So (for a symmetric tensor $\overleftrightarrow{T}$) we can always choose an orthonormal basis of eigenvectors.
$\hat{u}_i = \frac{\vec{v}_i}{|\vec{v}_i|}$ is also an eigenvector; call the eigenvalue T_i, so that

$$\overleftrightarrow{T} \cdot \hat{u}_i = T_i\hat{u}_i \tag{3.22}$$

And we can find the components in the basis $\{\hat{u}_i | i = 1, 2, 3\}$ as

$$\hat{u}_i \cdot \overleftrightarrow{T} \cdot \hat{u}_j = (\hat{u}_i \cdot \hat{u}_j)T_j = \delta_{ij}T_j \tag{3.23}$$

So

$$\overleftrightarrow{T} = T_1\hat{u}_1 \otimes \hat{u}_1 + T_2\hat{u}_2 \otimes \hat{u}_2 + T_3\hat{u}_3 \otimes \hat{u}_3 \tag{3.24}$$

This gives us the beginnings of a geometric interpretation of a tensor $\overleftrightarrow{T}$, at least if that tensor happens to be symmetric. A symmetric tensor defines a preferred set of orthogonal axes, with (not necessarily positive) numbers associated with each one.

There are different possibilities:

- All three eigenvalues $\{T_1, T_2, T_3\}$ are different. In that case the basis $\{\hat{u}_1, \hat{u}_2, \hat{u}_3\}$ is unique (up to cyclic reörderings of the basis vectors).
- Two eigenvalues are equal ($T_1 \neq T_2 = T_3$). Then $\hat{u}_1$ is unique (up to a sign), and $\hat{u}_2$ and $\hat{u}_3$ form an orthonormal basis for the plane perpendicular to $\hat{u}_1$.
- All three eigenvalues are equal ($T_1 = T_2 = T_3$). Then any vector is an eigenvector of $\overleftrightarrow{T}$ and the tensor is diagonal in any orthonormal basis. Symon calls this a "constant tensor", although a more appropriate term would be "isotropic". In this case, we can write $\overleftrightarrow{T} = T_1\overleftrightarrow{1}$ without reference to any basis.

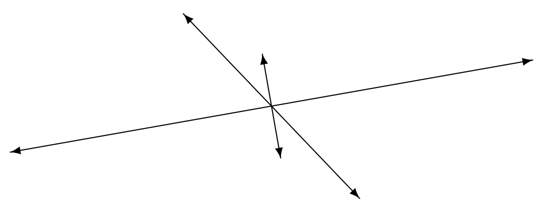
4 The Inertia Tensor

4.1 Review: Characterization of a Symmetric Tensor

Last time we saw *any* symmetric tensor $\overleftrightarrow{T}$ can be written

$$\overleftrightarrow{T} = T_1 \hat{u}_1 \otimes \hat{u}_1 + T_2 \hat{u}_2 \otimes \hat{u}_2 + T_3 \hat{u}_3 \otimes \hat{u}_3 \tag{4.1}$$

for some orthonormal basis $\{\hat{u}_i\}$ and set of eigenvalues $\{T_i\}$ which are "unique". So this gives us a physical picture of a symmetric tensor: three perpendicular axes with numbers attached. We could visualize it by drawing each axis as a double-headed arrow of the appropriate length:



(This doesn't work so well if any of the eigenvalues are negative.)

Different possibilities

- $\{T_1, T_2, T_3\}$ **all different**
  $\longrightarrow$ unit vectors $\{\hat{u}_1, \hat{u}_2, \hat{u}_3\}$ is unique (up to signs). In the case of the inertia tensor $\overleftrightarrow{I}$ these are called "principal axes of inertia"; they are usually associated with symmetries and often conveniently chosen as coördinate axes.

- $T_1 \neq T_2 = T_3$
  $\longrightarrow$ $\hat{u}_1$ unique (up to a sign); $\hat{u}_2$ and $\hat{u}_3$ can be any two perpendicular unit vectors in the plane perpendicular to $\hat{u}_1$.

- $T_1 = T_2 = T_3$
  $\overleftrightarrow{T} = T_1 \overleftrightarrow{1}$. Then the tensor is *isotropic* (Symon calls it "constant"). *Any* vector is an eigenvector, so any orthonormal basis diagonalizes $\overleftrightarrow{T}$.

4.2 Inertia Tensor in Detail

For a discrete mass distribution (collection of point masses)

$$\overleftrightarrow{I} = \sum_{k=1}^{N} m_k (\overleftrightarrow{1} r_k^2 - \vec{r}_k \otimes \vec{r}_k) \tag{4.2}$$

For a continuous mass distribution,

$$\overleftrightarrow{I} = \iiint \rho(\vec{r}) (\overleftrightarrow{1} r^2 - \vec{r} \otimes \vec{r}) \, d^3V \tag{4.3}$$

To examine the components in a particular basis, consider that

$$\begin{aligned}
\mathbf{1}r^2 - \mathbf{r}\mathbf{r}^{\mathrm{T}} &= \begin{pmatrix} x^2+y^2+z^2 & 0 & 0 \\ 0 & x^2+y^2+z^2 & 0 \\ 0 & 0 & x^2+y^2+z^2 \end{pmatrix} - \begin{pmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{pmatrix} \\
&= \begin{pmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{pmatrix}
\end{aligned} \tag{4.4}$$

so

$$I_{xx} = \iiint \rho(\vec{r})\,(y^2+z^2)\,d^3V \tag{4.5a}$$

$$I_{xy} = I_{yx} = -\iiint \rho(\vec{r})\,xy\,d^3V \tag{4.5b}$$

etc.

4.3 Example: Ellipsoid of Constant Density

As an example of how to calculate the components (4.5) explicitly, consider an ellipsoid $\mathcal{E}$ defined by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \le 1 \tag{4.6}$$

with constant density ρ. Now, Cartesian coördinates are not exactly suited to integrating over an ellipsoid, nor are standard spherical coördinates. But if we uses as our integration variables $\{\sigma, \vartheta, \varphi\}$ defined by

$$x = a\sigma\sin\vartheta\cos\varphi \tag{4.7a}$$

$$y = b\sigma\sin\vartheta\sin\varphi \tag{4.7b}$$

$$z = c\sigma\cos\vartheta \tag{4.7c}$$

Then the equation (4.6) defining the ellipsoid becomes just

$$\sigma^2 \le 1 \tag{4.8}$$

and to cover the whole ellipsoid, we just need the limits of integration

$$0 \le \sigma \le 1 \tag{4.9a}$$

$$0 \le \vartheta \le \pi \tag{4.9b}$$

$$0 \le \varphi < 2\pi \tag{4.9c}$$

The volume element is easy to see by analogy to spherical coördinates:

$$d^3V = dx\,dy\,dz\, = abc\,\sigma^2\,\sin\vartheta\,d\sigma\,d\vartheta\,d\varphi \tag{4.10}$$

It'll turn out that the components of the inertia tensor are more easily written in terms of the ellipsoid's mass than its density, so first we calculate the mass:

$$M = \rho abc \underbrace{\int_0^{2\pi} d\varphi}_{2\pi} \underbrace{\int_0^{\pi} \sin\vartheta\,d\vartheta}_{\int_{-1}^{1} d\mu = 2} \underbrace{\int_0^{1} \sigma^2\,d\sigma}_{\frac{1}{3}} = \frac{4\pi abc}{3}\rho \tag{4.11}$$

It's not hard to show that the off-diagonal components of the inertia tensor vanish in this case, since

$$\iiint\limits_{\mathcal{E}} xy\,dx\,dy\,dz = \iiint\limits_{\mathcal{E}} yz\,dx\,dy\,dz = \iiint\limits_{\mathcal{E}} xz\,dx\,dy\,dz = 0 \tag{4.12}$$

This is because the region of integration is symmetric under inversion $x \to -x$. This means that the integral of any quantity odd in x over this region vanishes, which takes care of xy and xz. Similarly, it's also symmetric under $y \to -y$, which means the integral of a function odd in y, such as yz, over the region vanishes as well.

To get the diagonal components, we need $\iiint\limits_{\mathcal{E}} x^2\,dx\,dy\,dz$, $\iiint\limits_{\mathcal{E}} y^2\,dx\,dy\,dz$, and $\iiint\limits_{\mathcal{E}} z^2\,dx\,dy\,dz$. The last is easiest to calculate explicitly in these coördinates:

$$\iiint\limits_{\mathcal{E}} z^2\,dx\,dy\,dz = abc^3 \underbrace{\int_0^{2\pi} d\varphi}_{2\pi} \underbrace{\int_0^{\pi} \cos^2\vartheta \sin\vartheta\,d\vartheta}_{\int_{-1}^{1}\mu^2 d\mu = \frac{2}{3}} \underbrace{\int_0^1 \sigma^4\,d\sigma}_{\frac{1}{5}} = \frac{4\pi abc}{15}c^2 = \frac{1}{\rho}\frac{M}{5}c^2 \tag{4.13}$$

The other two integrals can be calculated explicitly or by analogy, since everything in the problem is the same if we permute $x \to y \to z \to x$ and, $a \to b \to c \to a$:

$$\iiint\limits_{\mathcal{E}} x^2\,dx\,dy\,dz = \frac{1}{\rho}\frac{M}{5}a^2 \tag{4.14a}$$

$$\iiint\limits_{\mathcal{E}} y^2\,dx\,dy\,dz = \frac{1}{\rho}\frac{M}{5}b^2 \tag{4.14b}$$

We can now evaluate all the necessary integrals:

$$I_{xx} = \iiint\limits_{\mathcal{E}} \rho(y^2+z^2)\,dx\,dy\,dz = \frac{M}{5}(b^2+c^2) \tag{4.15a}$$

$$I_{yy} = \iiint\limits_{\mathcal{E}} \rho(x^2+z^2)\,dx\,dy\,dz = \frac{M}{5}(a^2+c^2) \tag{4.15b}$$

$$I_{zz} = \iiint\limits_{\mathcal{E}} \rho(x^2+y^2)\,dx\,dy\,dz = \frac{M}{5}(a^2+b^2) \tag{4.15c}$$

so, using this and the vanishing of the diagonal elements as explained above,

$$\begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} = \begin{pmatrix} \frac{M}{5}(b^2+c^2) & 0 & 0 \\ 0 & \frac{M}{5}(a^2+c^2) & 0 \\ 0 & 0 & \frac{M}{5}(a^2+b^2) \end{pmatrix} \tag{4.16}$$

4.4 Notes on the Inertia Tensor

Inertia tensor

$$\overleftrightarrow{I} = \iiint (\overleftrightarrow{1} r^2 - \vec{r}\otimes\vec{r})\,\rho(\vec{r})\,d^3V \tag{4.17}$$

Note, if a rigid body rotates in space, the distribution of mass will change, so its inertia tensor will change, but in a rotating reference frame which moves with the body, the components will stay the same. What's happening is that the body axes are changing but the moments of inertia about those axes are remaining the same. If the $\{\vec{e}_i'\}$ basis rotates with the body,

then

$$\overleftrightarrow{I} = \sum_{i=1}^{3} I_i \hat{u}_i \otimes \hat{u}_i = \sum_{i=1}^{3} I_{ii}' \vec{e}_i' \otimes \vec{e}_i' \tag{4.18}$$

so $I_{ij}' = \delta_{ij} I_i$ in this special basis.

4.4.1 Inertia Tensor Relative to Different Origins

Also note that because $\overleftrightarrow{I}$ is built from the position vector $\vec{r}$ it depends on the choice of origin.

$\vec{r}$ points to some position in the body; it's integrated over.

In particular, we can compare $\overleftrightarrow{I_{\mathcal{O}}}$ measured relative to some origin $\mathcal{O}$ with $\overleftrightarrow{I_{\mathcal{G}}}$ measured relevant to the center of mass $\mathcal{G}$. The position vectors relative to the two origins are related by

$$\vec{r}_{\mathcal{O}} = \vec{r}_{\mathcal{G}} - \vec{R} \tag{4.19}$$

where $\vec{R}$ is the center of mass position vector (relative to $\mathcal{O}$). For compactness of notation, we refer to $\vec{r}_{\mathcal{O}}$ as $\vec{r}$ and $\vec{r}_{\mathcal{G}}$ as $\vec{r}\,'$ so

$$\vec{r}\,' = \vec{r} - \vec{R} \tag{4.20}$$

The relationship between the two inertia tensors is simplified somewhat by the fact that since

$$\vec{R} = \frac{\iiint \vec{r}\,\rho(\vec{r})\,d^3V}{M} \tag{4.21}$$

we have

$$\iiint \vec{r}\,'\,\rho(\vec{r})\,d^3V = \iiint \vec{r}\,\rho(\vec{r})\,d^3V - \vec{R}\underbrace{\iiint \rho(\vec{r})\,d^3V}_{M} = M\vec{R} - M\vec{R} = \vec{0} \tag{4.22}$$

Now,

$$\begin{aligned} \overleftrightarrow{I_{\mathcal{O}}} &= \iiint \left[\overleftrightarrow{1}(\vec{r}\cdot\vec{r}) - \vec{r}\otimes\vec{r}\right]\rho(\vec{r})\,d^3V \\ &= \iiint \left[\overleftrightarrow{1}(\vec{R}+\vec{r}\,')\cdot(\vec{R}+\vec{r}\,') - (\vec{R}+\vec{r}\,')\otimes(\vec{R}+\vec{r}\,')\right]\rho(\vec{r})\,d^3V \\ &= \iiint \left[(\vec{R}\cdot\vec{R})\overleftrightarrow{1} - \vec{R}\otimes\vec{R}\right]\rho(\vec{r})\,d^3V + \iiint \left[(\vec{R}\cdot\vec{r}\,')\overleftrightarrow{1} - \vec{R}\otimes\vec{r}\,'\right]\rho(\vec{r})\,d^3V \\ &\quad + \iiint \left[(\vec{r}\,'\cdot\vec{R})\overleftrightarrow{1} - \vec{r}\,'\otimes\vec{R}\right]\rho(\vec{r})\,d^3V + \iiint \left[(\vec{r}\,'\cdot\vec{r}\,')\overleftrightarrow{1} - \vec{r}\,'\otimes\vec{r}\,'\right]\rho(\vec{r})\,d^3V \end{aligned} \tag{4.23}$$

In the first term the constant vector $\vec{R}$ can be pulled out of the integral to give

$$\iiint \left[(\vec{R}\cdot\vec{R})\overleftrightarrow{1} - \vec{R}\otimes\vec{R}\right]\rho(\vec{r})\,d^3V = \left[(\vec{R}\cdot\vec{R})\overleftrightarrow{1} - \vec{R}\otimes\vec{R}\right]\underbrace{\iiint \rho(\vec{r})\,d^3V}_{M} \tag{4.24}$$

The cross terms both vanish thanks to (4.22); for example

$$\iiint \left[(\vec{R}\cdot\vec{r}\,')\overleftrightarrow{1} - \vec{R}\otimes\vec{r}\,'\right]\rho(\vec{r})\,d^3V = \vec{R}\cdot\Big[\cancelto{\vec{0}}{\iiint \vec{r}\,'\rho(\vec{r})\,d^3V}\Big]\overleftrightarrow{1} - \vec{R}\otimes\Big[\cancelto{\vec{0}}{\iiint \vec{r}\,'\rho(\vec{r})\,d^3V}\Big] = \vec{0} \tag{4.25}$$

The last term is just $\overleftrightarrow{I_{\mathcal{G}}}$, since integrating $\vec{r}$ over the whole solid accomplishes the same think as integrating $\vec{r}\,'$ over the whole solid.

So in the end

$$\overleftrightarrow{I_{\mathcal{O}}} = \overleftrightarrow{I_{\mathcal{G}}} + M\left[(\vec{R}\cdot\vec{R})\overleftrightarrow{1} - \vec{R}\otimes\vec{R}\right] \tag{4.26}$$

I.e., a body's inertia tensor with respect to any origin is its inertia tensor with respect to its center of mass, plus the inertia tensor with respect to the origin of a pointlike object of a mass equal to the body's total mass located at the body's center of mass.

A Appendix: Correspondence to Class Lectures

Date	Sections	Pages	Topics
2004 March 30	1–2.3	2–5	Tensors: Derivation
2004 April 1	2.3.1–2.4	5–6	Tensors: More properties
2004 April 20	3	6–9	Tensors: Still more properties
2004 April 22	4.1–4.3	10–12	The Inertia Tensor
2004 April 27	4.4–4.4.1	12–14	Notes on the Inertia Tensor