1 The inertia tensor

P. V. Buividovich

1.1 What one should already know about rotating planar (or cylindrically symmetric) systems

Angular momentum is related to angular velocity by

\[ L = Iw \]  \hspace{1cm} (1)

Kinetic energy is related to angular velocity by

\[ K = \frac{Iw^2}{2} \]  \hspace{1cm} (2)

To refresh our knowledge, put your bets on the following race: there is an inclined plane and four round objects on it, all of the same mass and radius:

- A disc
- A ring
- A ball (uniformly filled inside)
- A hollow sphere

Assume infinite friction, so no sliding, only rolling (and no dissipation of energy into heat). Which object will roll down faster and why?

1.2 Angular momentum of a rigid 3D body

Now, however, we are interested in a general 3D rotation of a non-planar rigid bodies, in general without any symmetry! First consider angular momentum of a system of point masses:

\[ \vec{L} = \sum_{\alpha} \vec{r}_\alpha \times \vec{p}_\alpha, \]  \hspace{1cm} (3)

where \( \alpha \) labels different point masses with positions \( \vec{r}_\alpha \) and momenta \( \vec{p}_\alpha \). Assuming that all the masses are rotating with angular velocity \( \vec{w} \), we can express the momenta \( \vec{p}_\alpha \) as

\[ \vec{p}_\alpha = m_\alpha \vec{w} \times \vec{r}_\alpha. \]  \hspace{1cm} (4)
Combining (3) and (4), and using the vector identity $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$ we get

$$L = \sum_\alpha m_\alpha \vec{r}_\alpha \times (\vec{w} \times \vec{r}_\alpha) = \sum_\alpha \left( m_\alpha \vec{r}_\alpha^2 \vec{w} - m_\alpha \vec{r}_\alpha (\vec{r}_\alpha \cdot \vec{w}) \right)$$

(5)

It is important to notice that for a general rigid 3D system of masses, the angular momentum vector is not necessarily parallel to the angular velocity vector! We’ll discuss some of the interesting consequences of this fact, such as the precession of the top, at the end of this lecture.

If we now assume that we have a planar 2D system with $\vec{w}$ being perpendicular to its plane, so that $(\vec{r}_\alpha \cdot \vec{w}) = 0$ we immediately find that $L = \sum_\alpha m_\alpha \vec{r}_\alpha^2 \vec{w}$, where $I = \sum_\alpha m_\alpha \vec{r}_\alpha^2$ is the moment of inertia which was studied in the first year mechanics course.

Let us now write this relation in component-wise form, using some basis where $\vec{r} = (x, y, z)$:

$$
\begin{pmatrix}
L_x \\
L_y \\
L_z
\end{pmatrix} =
\left( \sum_\alpha \begin{pmatrix}
  m_\alpha (y_\alpha^2 + z_\alpha^2) & -m_\alpha x_\alpha y_\alpha & -m_\alpha x_\alpha z_\alpha \\
  -m_\alpha x_\alpha y_\alpha & m_\alpha (x_\alpha^2 + z_\alpha^2) & -m_\alpha y_\alpha z_\alpha \\
  -m_\alpha x_\alpha z_\alpha & -m_\alpha y_\alpha z_\alpha & m_\alpha (x_\alpha^2 + y_\alpha^2)
\end{pmatrix}
\right)
\begin{pmatrix}
w_x \\
w_y \\
w_z
\end{pmatrix}
$$

(6)

where we have introduced the $3 \times 3$ symmetric inertia matrix $I$, which has in total 6 independent components. In compact form this matrix can be written as

$$I_{ij} = \sum_\alpha m_\alpha \left( r_\alpha^2 \delta_{ij} - r_\alpha i r_\alpha j \right).$$

(7)

In vector-tensor notation,

$$I = \sum_\alpha m_\alpha \left( \vec{r}_\alpha^2 \mathbb{1} - \vec{r}_\alpha \otimes \vec{r}_\alpha \right),$$

(8)

where $\mathbb{1}$ denotes the $3 \times 3$ identity matrix.

If we have a continuous 3D rigid body, we can split it into infinitesimal volumes $d^3\vec{r} = dr_1 dr_2 dr_3$ with masses $dm(\vec{r}) = \rho(\vec{r}) d^3\vec{r}$, where $\rho(\vec{r})$ is the
mass density. The sums in the above definitions are then replaced by 3D integrals over the volume $V$ of our rigid body:

$$I_{ij} = \int_V d^3\vec{r} \rho(\vec{r}) \left(r^2 \delta_{ij} - r_i r_j\right).$$

(9)

Or, in vector-tensor notation

$$I = \int_V d^3\vec{r} \rho(\vec{r}) \left(r^2 \mathbf{1} - \vec{r} \otimes \vec{r}\right).$$

(10)

### 1.3 Kinetic energy of a rigid 3D body

Let us now calculate the kinetic energy of our system of point masses:

$$K = \sum_{\alpha} \frac{m_{\alpha} \vec{v}_{\alpha}^2}{2} = \sum_{\alpha} \frac{m_{\alpha} (\vec{r}_{\alpha} \times \vec{w})^2}{2}.\quad (11)$$

We now use another vector identity $(\vec{a} \times \vec{b})^2 = \vec{a}^2 \vec{b}^2 - (\vec{a} \cdot \vec{b})^2$ to rewrite this as

$$K = \frac{1}{2} \sum_{\alpha} m_{\alpha} \left(\vec{r}_{\alpha}^2 \vec{w}^2 - (\vec{r}_{\alpha} \vec{w})^2\right).\quad (12)$$

Again using some basis where $\vec{r}$ has components $\vec{r} = (x, y, z)$, we can rewrite this formula as

$$K = \frac{1}{2} \left(\sum_{\alpha} m_{\alpha} \left(y_{\alpha}^2 + z_{\alpha}^2\right) w_x^2 + m_{\alpha} \left(x_{\alpha}^2 + z_{\alpha}^2\right) w_y^2 + m_{\alpha} \left(x_{\alpha}^2 + y_{\alpha}^2\right) w_z^2 - 2m_{\alpha} x_{\alpha} y_{\alpha} w_x w_y - 2m_{\alpha} x_{\alpha} z_{\alpha} w_x w_z - 2m_{\alpha} y_{\alpha} z_{\alpha} w_y w_z\right).\quad (13)$$

Having a careful look at the coefficients of the terms like $w_x^2$ or $w_x w_y$, we immediately recognize that the coefficients of these terms are proportional the components of the inertia matrix introduced in (6), and we can write the kinetic energy as

$$K = \frac{1}{2} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}.\quad (14)$$

We again see that the kinetic energy can be written in a very compact form using the inertia matrix. Again, recall the simpler formula $K = \frac{1}{2} I \vec{w}^2$!
1.4 Inertia matrix as a tensor

The choice of the basis in which the inertia tensor is calculated is arbitrary. If we have calculated it in some special basis, do we have to repeat our
calculation for another coordinate system? No! The reason is that the inertia
matrix transforms as a tensor under rotations of the coordinate frame. Let
us consider another coordinate system \((x', y', z')\), related to our previous
coordinates \((x, y, z)\) as

\[
\begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix} = \begin{pmatrix}
\Omega_{xx} & \Omega_{xy} & \Omega_{xz} \\
\Omega_{xy} & \Omega_{yy} & \Omega_{yz} \\
\Omega_{xz} & \Omega_{yz} & \Omega_{zz}
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix},
\]

(15)
or

\[
r'_i = \sum_j \Omega_{i,j} r_j
\]

(16)

Requiring that the length of any vector remains the same in new coor-
dinate system, we get \(\Omega \Omega^T = 1\), or \(\Omega^T = \Omega^{-1}\) (where \(-1\) denotes matrix
inversion). Such matrices are called orthogonal.

So, let’s check how does the matrix \(I\) transform. The expression (7) in
the new coordinates is

\[
I'_{ij} = \sum_\alpha m_\alpha \left( r'^2_\alpha \delta_{ij} - \sum_{k,l} \Omega_{ik} \Omega_{jl} r'_k r'_l \right),
\]

(17)

where we have used the fact that rotations of the coordinate frame do not
change the length of vectors \(r'_\alpha\). Using the identity \(1 = \Omega \Omega^T\) which can also
be written as \(\delta_{ij} = \sum_{k,l} \Omega_{ik} \Omega_{jl} \delta_{ij}\) we can write

\[
I'_{ij} = \Omega_{ik} \Omega_{jl} I_{kl} = \Omega_{ik} I_{kl} \left(\Omega^T\right)_{lj},
\]

(18)
or, in compact matrix notation

\[
I' = \Omega I \Omega^T = \Omega I \Omega^{-1},
\]

(19)

where we have used the orthogonality of the matrix \(\Omega\). The objects which
transform in such a way under the change of the coordinate system are called
tensors. In general, any coefficient relating two vectors is, strictly speaking,
a tensor. Some other physical examples of tensors are the dielectric permi-
tivity and conductivity.

The above relations (18) and (19) are also very important in the general
case of the rotation of an asymmetric rigid body - they show that as the
body rotates, the tensor of inertia also changes! This is in fact the origin of
the precession phenomenon.

1.5 Principal moments of inertia and principal axes

Recall that the rotation of the coordinate system can be specified by three real
numbers - e.g. Euler angles. For this reason one can choose the coordinate
system in such a way that 6 components of the inertia matrix $I_{ij}$ reduce down
to 3 components. The most natural choice is to make the matrix $I$ diagonal -
such that for rotations with angular velocity $\vec{\omega} = w_i \vec{e}_i'$ around each particular
coordinate axis $i$ (with basis vector $\vec{e}_i'$) we have $\vec{L} = I_i w_i \vec{e}_i'$. Let us denote the
components of angular velocity vector $\vec{\omega} = w_i \vec{e}_i'$ in our original coordinate
system as $(w_x, w_y, w_z)$. Then we have the following equation defining the
rotation around one of the principal axes of rotation:

$$
\begin{pmatrix}
L_x \\
L_y \\
L_z 
\end{pmatrix} = 
\begin{pmatrix}
I_{xx} & I_{xy} & I_{xz} \\
I_{xy} & I_{yy} & I_{yz} \\
I_{xz} & I_{yz} & I_{zz} 
\end{pmatrix} 
\begin{pmatrix}
w_x \\
w_y \\
w_z 
\end{pmatrix} = 
I 
\begin{pmatrix}
w_x \\
w_y \\
w_z 
\end{pmatrix},
$$

(20)

or, in other words,

$$
\begin{pmatrix}
(I_{xx} - I) & I_{xy} & I_{xz} \\
I_{xy} & (I_{yy} - I) & I_{yz} \\
I_{xz} & I_{yz} & (I_{zz} - I) 
\end{pmatrix} 
\begin{pmatrix}
w_x \\
w_y \\
w_z 
\end{pmatrix} = 0.
$$

(21)

As you should know from the linear algebra course, a necessary condition for
this equation to have solutions is

$$
\det 
\begin{pmatrix}
(I_{xx} - I) & I_{xy} & I_{xz} \\
I_{xy} & (I_{yy} - I) & I_{yz} \\
I_{xz} & I_{yz} & (I_{zz} - I) 
\end{pmatrix} = 0.
$$

(22)

This is a cubic equation, which have three real solutions $I = I_1, I = I_2, I = I_3$
- which are called the principal moments of inertia, or eigenvalues of the in-
ertia matrix. In general, these solutions can be quite complicated, as the
general formula for the solution of the cubic equations is quite complicated.
Once we’ve found these solutions, we can plug each of them back into equation (21) and find the corresponding ratios of the components $w_x : w_y : w_z$. Note that since (21) is a homogeneous linear equation, the absolute value of $\bar{w} = (w_x, w_y, w_z)$ is not fixed. For each eigenvalue $I_1$, $I_2$, $I_3$ we have a different direction of $\bar{w}$, which we can parameterize by unit vectors $\vec{e}_1', \vec{e}_2', \vec{e}_3'$ - the eigenvectors of the inertia matrix $I$. One can easily prove that these eigenvectors are mutually orthogonal - this is a general property of eigenvectors of real symmetric matrices. Let me also mention that the notion of eigenvectors/eigenvalues will be central to the quantum mechanics course.

At the practical level, knowing the principal axes of rotation is extremely useful, since one needs to calculate only 3 numbers instead of 6. In practice, if a rigid body has reflection or rotation symmetry, the principal axes will be aligned with either with the reflection plane or with rotation axis.

### 1.6 Transformations under shifts

Yet another redundancy in the choice of the coordinate system in which we calculate the inertia tensor is the shift of the origin of the coordinate system. Let us check how the inertia tensor transform under a global shift $\vec{r}' = \vec{r} + \vec{R}$:

$$ I' = \sum_\alpha m_\alpha \left( (\vec{r}')^2 \mathbb{1} - \vec{r}' \otimes \vec{r}' \right) = \sum_\alpha m_\alpha \left( (\vec{r}_\alpha^2 + 2\vec{r}_\alpha \vec{R} + \vec{R}^2) \mathbb{1} - \vec{r}_\alpha \otimes \vec{r}_\alpha - \vec{r}_\alpha \otimes \vec{R} - \vec{R} \otimes \vec{r}_\alpha - \vec{R} \otimes \vec{R} \right) $$

Let us now introduce the total mass $M$ of the system as well as the vector of distance $\vec{\rho}$ to the center of mass of the rigid body:

$$ M = \sum_\alpha m_\alpha, \quad \vec{\rho} = \frac{1}{M} \sum_\alpha m_\alpha \vec{r}_\alpha. $$

Then we can write $I'$ in the following compact form:

$$ I' = I + I_{\text{point}}(M, \vec{R}) + 2M (\vec{\rho} \cdot \vec{R}) \mathbb{1} - M \vec{\rho} \otimes \vec{R} - M \vec{R} \otimes \vec{\rho}, $$

where $I_{\text{point}}(M, \vec{R})$ is the moment of inertia of a point mass $M$ at position $\vec{R}$:

$$ I_{\text{point}}(M, \vec{R}) = M \vec{R}^2 \mathbb{1} - M \vec{R} \otimes \vec{R}. $$
We now see that when calculating the tensor of inertia, it is most advantageous to set the origin of the coordinate system at the center of mass of the rigid body, so that $\vec{\rho} = 0$ - in this case the inertia tensor with respect to any other point has a particularly simple form (which is sometimes referred to as the parallel axis theorem).

1.7 Example: inertia tensor of a brick

To illustrate all of the above, let’s calculate the inertia tensor for some body which has no cylindric or spherical symmetry - for example, the uniform brick! Let the sides of the brick be $a, b, c$, the mass of the brick be $M = abc\rho$, where $\rho$ is the density of the material of the brick. Since the brick has three reflection planes, it is natural to choose the center of the brick as the origin, and the coordinate axes perpendicular to the surfaces of the brick. It is easy to show that indeed in such a coordinate system only diagonal components of the inertia tensor will be nonzero. Let’s now calculate, for instance, $I_{zz}$:

$$I_{zz} = \rho \int_{-c/2}^{c/2} dz \int_{-b/2}^{b/2} dy \int_{-a/2}^{a/2} dx \left( x^2 + y^2 \right) =$$

$$= \rho \int_{-c/2}^{c/2} dz \int_{-b/2}^{b/2} dy \left( \frac{a^3}{12} + ay^2 \right) =$$

$$= \rho \int_{-c/2}^{c/2} dz \frac{a^3 b + ab^3}{12} = \rho \frac{a^3 b c + b a c^3}{12} = M a^2 + b^2 \quad (27)$$

The other components can be found by simply permuting the widths $a, b, c$ in the above expression. We find, then

$$I_{\text{brick}} = \begin{pmatrix}
M \frac{b^2 + c^2}{12} & 0 & 0 \\
0 & M \frac{a^2 + c^2}{12} & 0 \\
0 & 0 & M \frac{a^2 + b^2}{12}
\end{pmatrix} \quad (28)$$

1.8 Example: precession

Consider now the free rotation of some rigid body with principal moments of inertia $I_1, I_2, I_3$. The principal axes of the inertia tensor can be used to
define the coordinate system and the reference frame. However, since the body rotates, this reference frame is, in general, not inertial! For a rotation around the center of mass we can in general write

$$\left( \frac{d\vec{L}}{dt} \right)_{\text{inertial}} = \left( \frac{d\vec{L}}{dt} \right)_{\text{body}} + \vec{\omega} \times \vec{L} = 0. \quad (29)$$

Writing these equations in the reference frame associated with the body, we get the Euler’s equations:

$$I_1 \dot{w}_1 = (I_2 - I_3) w_2 w_3$$
$$I_2 \dot{w}_2 = (I_3 - I_1) w_1 w_3$$
$$I_3 \dot{w}_3 = (I_1 - I_2) w_1 w_2 \quad (30)$$

In general, there is no analytic solution for these equations! Assume, however, cylindric symmetry around some axis, so that $I_1 = I_2 = I$ and $\Delta I = I_3 - I_1 = I_3 - I_2 \neq 0$. From the last equation it immediately follows that $w_3 = \text{const}$. The first two equations reduce to

$$\dot{w}_1 = -\frac{\Delta I}{I} w_2 w_3$$
$$\dot{w}_2 = \frac{\Delta I}{I} w_1 w_3, \quad (31)$$

or

$$\ddot{w}_1 = -\left( \frac{\Delta I}{I} w_3 \right)^2 w_1, \quad (32)$$

which is nothing but the equation for harmonic oscillations! This means that if initially the angular velocity is not parallel with the symmetry axis and the angular momentum, the body slowly precesses with angular frequency $\omega = \frac{\Delta I}{I} w_3$.

### 1.9 Interesting example for consideration at home

An interesting problem with somewhat counterintuitive result (Dzhanibekov effect) is the stability of the Euler equations (30). Try to apply Lyapunov’s criteria to identify when the free rotation becomes unstable.