1 The inertia tensor

1.1 Kinetic energy

Remaining consistent with the approach we followed so far in the course, we describe a rigid body as made of a large number of point masses m_i with positions $\mathbf{r}_i$ and constraints among them. As this body evolves in space under the effect of external forces, it is a sensible idea to formally separate the motion of the center of mass of the object (translational motion) and the motion of the body around its center of mass (rotational motion). This is done mathematically as follows.

Let us first define the position of the center of mass by generalizing the definition we had used for the 2-body system in Lecture 2:

$$\mathbf{R} = \frac{1}{M}\sum_i m_i \mathbf{r}_i \tag{1}$$

where $M = \sum_i m_i$ is the total mass of the body.

We then introduce the vectors $\mathbf{x}_i$ from the center of mass to the point pass m_i of the body:

$$\mathbf{r}_i = \mathbf{R} + \mathbf{x}_i$$

It turns out that the total kinetic energy of the rigid body takes a fairly simple form in terms of $\mathbf{R}$ and $\mathbf{x}_i$. Indeed, we have

$$\begin{aligned} E_K &= \frac{1}{2}\sum_i m_i|\dot{\mathbf{r}}_i|^2 = \frac{1}{2}\sum_i m_i(\dot{\mathbf{R}} + \dot{\mathbf{x}}_i)\cdot(\dot{\mathbf{R}} + \dot{\mathbf{x}}_i) \\ &= \frac{1}{2}\sum_i m_i\left(|\dot{\mathbf{R}}|^2 + 2\dot{\mathbf{x}}_i\cdot\dot{\mathbf{R}} + |\dot{\mathbf{x}}_i|^2\right) \\ &= \frac{1}{2}M|\dot{\mathbf{R}}|^2 + \frac{1}{2}\sum_i m_i\left(2\dot{\mathbf{x}}_i\cdot\dot{\mathbf{R}} + |\dot{\mathbf{x}}_i|^2\right) \end{aligned}$$

Now, note that

$$\sum_i m_i\dot{\mathbf{x}}_i\cdot\dot{\mathbf{R}} = \dot{\mathbf{R}}\cdot\sum_i m_i\dot{\mathbf{x}}_i = \dot{\mathbf{R}}\cdot\sum_i m_i(\dot{\mathbf{r}}_i - \dot{\mathbf{R}}) = \dot{\mathbf{R}}\cdot(M\dot{\mathbf{R}} - M\dot{\mathbf{R}}) = 0$$

We conclude that the kinetic energy can be written as

$$E_K = E_C + E_R = \frac{1}{2}M|\dot{\mathbf{R}}|^2 + \frac{1}{2}\sum_i m_i|\dot{\mathbf{x}}_i|^2 \tag{2}$$

In other words, the kinetic energy is the sum of the kinetic energy E_C of the translational motion of the total mass at the center of mass and the kinetic energy E_R of rotation around the center of mass.

1.2 Inertia tensor

Euler's rotation theorem states that in three-dimensional space, any displacement of a rigid body such that a point on the rigid body remains fixed, is equivalent to a single rotation about some axis that runs through the fixed point. For our purposes, that means that we can always find a rotation axis and a rotation angle to represent the change of all the $\mathbf{x}_i$ from a time t_1 to a time t_2 as a rotation (since the center of mass $\mathbf{R}$ is fixed in the frame moving at the velocity $\dot{\mathbf{R}}$). Taking the limit $\Delta t = t_2 - t_1 \to 0$, this rotation becomes an instantaneous rotation, which can be represented with the angular velocity vector $\mathbf{\Omega}$. $\mathbf{\Omega}$ points in the direction of the axis of rotation, and its magnitude is equal to the rate of rotation.

In terms of $\mathbf{\Omega}$, the evolution of the $\mathbf{x}_i$ in time takes the particularly simple form

$$\dot{\mathbf{x}}_i = \mathbf{\Omega}\times\mathbf{x}_i \tag{3}$$

We are now ready to express the kinetic energy of rotation E_R in terms of $\mathbf{\Omega}$ and the $\mathbf{x}_i$:

$$E_R = \frac{1}{2}\sum_i m_i\left(\mathbf{\Omega}\times\mathbf{x}_i\right)\cdot\left(\mathbf{\Omega}\times\mathbf{x}_i\right)$$

Let us choose a rectangular coordinate system with arbitrary orientation, origin at the center of mass and basis of unit vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Writing $\mathbf{\Omega} = \Omega_1\mathbf{e}_1 + \Omega_2\mathbf{e}_2 + \Omega_3\mathbf{e}_3$, we can rewrite E_R as

$$\begin{aligned} E_R &= \frac{1}{2}\sum_i m_i \left(\sum_j \Omega_j \mathbf{e}_j \times \mathbf{x}_i\right) \cdot \left(\sum_k \Omega_k \mathbf{e}_k \times \mathbf{x}_i\right) \\ &= \frac{1}{2}\sum_{j,k} \Omega_j \Omega_k \sum_i m_i \left(\mathbf{e}_j \times \mathbf{x}_i\right) \cdot \left(\mathbf{e}_k \times \mathbf{x}_i\right) \\ &= \frac{1}{2}\mathbf{\Omega}^T \mathbf{I} \mathbf{\Omega} \end{aligned}$$

where the matrix $\mathbf{I}$ is called the *inertia tensor*, and has the following entries

$$I_{jk} = \sum_i m_i \left(\mathbf{e}_j \times \mathbf{x}_i\right) \cdot \left(\mathbf{e}_k \times \mathbf{x}_i\right) \tag{4}$$

The inertia tensor only depends on the distribution of mass in the body with respect to the coordinate system we have chosen. It is a quantity that is intrinsic to the body under consideration. What is more, through a somewhat clever representation of E_R, we have demonstrated that the rotational kinetic energy of a rigid body only depends on second order moments of the mass distribution. This is quite remarkable. One may have expected the rotational energy to depend on subtle details of the mass distribution, expressed through higher order moments.

2 Properties of the inertia tensor

Since the dot product commutes, it is clear that $\mathbf{I}$ is **symmetric**. This is the first, obvious property of $\mathbf{I}$, which will have important implications below. Let us investigate a few more important properties.

2.1 Change of basis

Since our choice of rectangular basis used in the previous section was arbitrary, it is instructive to see how the matrix $\mathbf{I}$ changes when it is expressed in an another basis. Let us consider the rotation matrix $\mathbf{R}$ that rotates the basis vectors (e_1, e_2, e_3) to the new basis vectors $(e_1^{'}, e_2^{'}, e_3^{'})$:

$$\mathbf{e}_j^{'} = \mathbf{R}\mathbf{e}_j \qquad j = 1, 2, 3$$

where $\mathbf{R}$ is a real orthogonal matrix. The components of $\mathbf{\Omega}$ in the rotated basis are obtained by doting $\mathbf{\Omega}$ with each basis vector:

$$\Omega_j^{'} = \mathbf{\Omega} \cdot \mathbf{e}_j^{'} = \mathbf{\Omega} \cdot \mathbf{R}\mathbf{e}_j \tag{5}$$

Now, the rotation of two vectors conserves the dot product between these two vectors. In other words, for two vectors $\mathbf{u}$ and $\mathbf{v}$ and a rotation $\mathbf{R}$ we can say

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{R}\mathbf{u} \cdot \mathbf{R}\mathbf{v}$$

Hence, for these two vectors, we can also write

$$\mathbf{R}^{-1}\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{R}\mathbf{v}$$

Applying this property to Equation (5), we find

$$\Omega_j^{'} = \mathbf{R}^{-1}\mathbf{\Omega} \cdot \mathbf{e}_j \tag{6}$$

In other words, the components of $\mathbf{\Omega}$ with respect to the rotated basis are the same as the components of $\mathbf{R}^{-1}\mathbf{\Omega}$ with respect to the original basis. We may write the short-handed relation

$$\mathbf{\Omega}^{'} = \mathbf{R}^{-1}\mathbf{\Omega} \quad \Leftrightarrow \quad \mathbf{\Omega} = \mathbf{R}\mathbf{\Omega}^{'} \tag{7}$$

where $\mathbf{\Omega}$ and $\mathbf{\Omega}^{'}$ are the components of the rotation $\mathbf{\Omega}$ in the basis $\mathbf{e}_j$ and $\mathbf{e}_j^{'}$ respectively.

We are now ready to compute the expression for the transformation of the inertia tensor associated with a change of basis. The rotational energy is

$$E_R = \frac{1}{2}\mathbf{\Omega}^T\mathbf{I}\mathbf{\Omega} = \frac{1}{2}\mathbf{\Omega}^{'T}\mathbf{R}^T\mathbf{I}\mathbf{R}\mathbf{\Omega}^{'} \tag{8}$$

E_R as expressed above is the kinetic energy of rotation expressed with respect to the $\mathbf{e}_j'$ basis. It must therefore be that the inertia tensor in this basis is given by

$$\mathbf{I}^{'} = \mathbf{R}^T\mathbf{I}\mathbf{R} \tag{9}$$

2.2 Moments of inertia

Let us go back to the expression (4) for $\mathbf{I}$ and write $\mathbf{I}$ explicitly by writing all the $\mathbf{x}_i$ in terms of their components in the basis $\mathbf{e}_j$: $\mathbf{x}_i = (a_i, b_i, c_i)$. Using the vector identity $(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C})$, the entries of the matrix $\mathbf{I}$ are readily computed, and one finds

$$\mathbf{I} = \begin{pmatrix} \sum_i m_i(b_i^2+c_i^2) & -\sum_i m_i a_i b_i & -\sum_i m_i a_i c_i \\ -\sum_i m_i a_i b_i & \sum_i m_i(a_i^2+c_i^2) & -\sum_i m_i b_i c_i \\ -\sum_i m_i a_i c_i & -\sum_i m_i b_i c_i & \sum_i m_i(a_i^2+b_i^2) \end{pmatrix}$$

We define the *moment of inertia* I about a given line by the expression

$$I = \sum_i m_i \left(d_i^{\perp}\right)^2$$

where $d_i^{\perp}$ is the perpendicular distance from the line to the point mass with index i. The diagonal entries of $\mathbf{I}$ can then be given the following physical interpretation: they are the moments of inertia about the lines coinciding with the coordinate axes associated with the basis $\mathbf{e}_j$. The off-diagonal components of $\mathbf{I}$ are called *products of inertia*.

2.3 Principal Moments of Inertia

A reasonable question to ask, in the light of formula (9) is whether it is possible to find a rotation $\mathbf{R}$ and a basis in which the inertia tensor $\mathbf{I}$ is diagonal. The answer is yes, since $\mathbf{I}$ is real and symmetric. Indeed, we know that for every real symmetric matrix $\mathbf{M}$ there exists a real orthogonal matrix $\mathbf{R}$ (in our case we called it a rotation) such that $\mathbf{D} = \mathbf{R}^T\mathbf{M}\mathbf{R}$ is a diagonal matrix

Consider the basis $\mathbf{e}_j^{'}$ in which the inertia tensor $\mathbf{I}'$ is diagonal. The axes in the body through the center of mass and aligned with the coordinate axes are called the *principal axes*. The diagonal entries of $\mathbf{I}^{'}$, its only nonzero entries, are called the *principal moments of inertia*. The rotational kinetic energy E_R has a particularly simple form when $\mathbf{I}'$ is expressed in the basis in which it is orthogonal. If $(\Omega_1, \Omega_2, \Omega_3)$ are the components of $\mathbf{\Omega}$ on the principal axes, then we can write

$$E_R = \frac{1}{2}\left[I_{11}^{'}\Omega_1^2 + I_{22}^{'}\Omega_2^2 + I_{33}^{'}\Omega_3^2\right] \tag{10}$$

3 Representing the angular momentum and the angular velocity vector

3.1 Angular momentum

In Lecture 2, we defined the angular momentum vector $\mathbf{L}$ of a point mass with mass m and position $\mathbf{r}$ orbiting around a fixed center:

$$\mathbf{L} = \mathbf{r}\times(m\dot{\mathbf{r}})$$

By analogy, we define the angular momentum $\mathbf{L}$ of the rigid body as

$$\mathbf{L} = \sum_i \mathbf{r}_i \times (m_i\dot{\mathbf{r}}_i)$$

Let us plug the transformation $\mathbf{r}_i = \mathbf{R} + \mathbf{x}_i$ in the expression above.

We have

$$\begin{aligned}\mathbf{L} &= \sum_i m_i \left(\mathbf{R} + \mathbf{x}_i\right) \times \left(\dot{\mathbf{R}} + \dot{\mathbf{x}}_i\right) \\ &= \mathbf{R} \times \left(M\dot{\mathbf{R}}\right) + \mathbf{R} \times \sum_i m_i \dot{\mathbf{x}}_i + \sum_i m_i \mathbf{x}_i \times \mathbf{R} + \sum_i \mathbf{x}_i \times (m_i \dot{\mathbf{x}}_i) \\ &= \mathbf{R} \times \left(M\dot{\mathbf{R}}\right) + \sum_i \mathbf{x}_i \times (m_i \dot{\mathbf{x}}_i)\end{aligned}$$

We conclude that

$$\mathbf{L} = \mathbf{L}_C + \mathbf{L}_R \tag{11}$$

with

$$\mathbf{L}_C = \mathbf{R} \times \left(M\dot{\mathbf{R}}\right) \quad \mathbf{L}_R = \sum_i \mathbf{x}_i \times (m_i \dot{\mathbf{x}}_i) \tag{12}$$

We just showed that the angular momentum can be naturally expressed in terms of the sum of the angular momentum $\mathbf{L}_C$ of the center of mass and the rotational angular momentum $\mathbf{L}_R$.

The rotational angular momentum has a very simple expression in terms of the inertia tensor $\mathbf{I}$. Writing $\dot{\mathbf{x}}_i = \mathbf{\Omega} \times \mathbf{x}_i$, and writing $\mathbf{\Omega}$ in terms of its components $(\Omega_1, \Omega_2, \Omega_3)$, one finds that each component j of $\mathbf{L}_R$ is given by

$$L_{R,j} = \sum_k I_{jk} \Omega_k$$

In other words,

$$\mathbf{L}_R = \mathbf{I}\mathbf{\Omega} \tag{13}$$

When $\mathbf{I}$ is written in a basis in which it is diagonal, the components of $\mathbf{L}$ have a particularly simple form:

$$L_{R,1} = I_{11}\Omega_1 \qquad L_{R,2} = I_{22}\Omega_2 \qquad L_{R,3} = I_{33}\Omega_3 \tag{14}$$

3.2 Components of the angular velocity vector

We have found convenient representations for the rotational kinetic energy and the angular momentum of rotation in terms of the inertia tensor and the angular velocity vector. We have seen how to calculate $\mathbf{I}$ in a basis $\mathbf{e}_j$, and we also so that there exists a basis in which $\mathbf{I}$ is diagonal. Let us now see how we would calculate the components $\mathbf{\Omega}' = (\Omega_1, \Omega_2, \Omega_3)$ of $\mathbf{\Omega}$ in that basis.

Let us consider generalized coordinates $\mathbf{q}(t)$ that we use to describe the evolution of the orientation of the body under study as it moves. The natural way to specify the orientation of the body is to take a reference orientation with respect to some fixed basis, and to introduce the rotation $\mathbf{R}(\mathbf{q}(t))$ that takes the body from the reference orientation to the current orientation. A natural idea is to take the reference orientation such that the principal axes of the body are aligned with the basis $\mathbf{e}_j$.

Now let $\mathbf{x}_i(t)$ be the positions of the point particles in the body at some time t, and $\mathbf{x}_i'$ be the positions of the point particles in the reference orientation. Mathematically, the rotation $\mathbf{R}(\mathbf{q}(t))$ is such that

$$\mathbf{x}_i(t) = \mathbf{R}(\mathbf{q}(t))\mathbf{x}_i'$$

Our goal is see how the $\mathbf{x}_i(t)$ evolve in time according to the relation above, and connect this evolution to the evolution equation (3) to obtain a formula for $\mathbf{\Omega}$.

We have

$$\dot{\mathbf{x}}_i(t) = \frac{d\mathbf{R}(\mathbf{q}(t))}{dt}\mathbf{x}_i' = \frac{d\mathbf{R}}{dt}\mathbf{R}^T\mathbf{x}_i(t) \tag{15}$$

where we have used the fact that $\mathbf{R}$, a rotation matrix, is orthogonal. Comparing Equations (3) and (15), we can make the formal identification

$$\mathbf{\Omega}\times \qquad \leftrightarrow \qquad \frac{d\mathbf{R}}{dt}\mathbf{R}^T$$

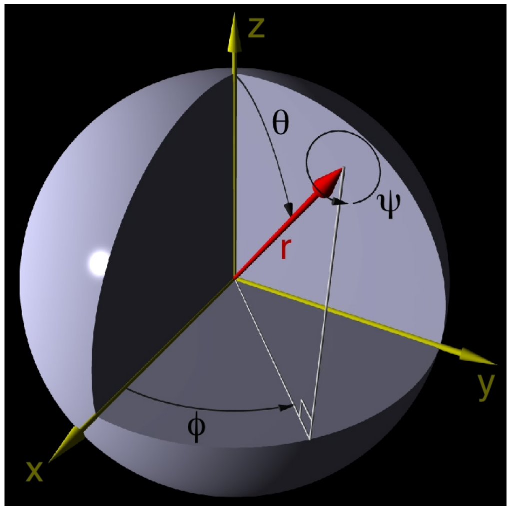

Figure 1: Illustration of the Euler angles with our convention

Now, we know that for any vector $\mathbf{u} = (u_1, u_2, u_3)$, the matrix operator $\mathbf{C}(\mathbf{u})$ corresponding to the operator $\mathbf{u}\times$ is given by the following skew-symmetric matrix:

$$\mathbf{C} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$$

Conversely, from any skew-symmetric matrix $\mathbf{M}$, one can easily reconstruct the vector $\mathbf{u}$ such that $\mathbf{M} = \mathbf{u}\times$, simply by extracting the bottom off-diagonal entries and ordering them appropriately. Let us call this operation $\mathbf{C}^{-1}$. The result of our analysis thus far is that the components of the angular velocity vector in the fixed basis are given by

$$\mathbf{\Omega} = \mathbf{C}^{-1}\left(\frac{d\mathbf{R}}{dt}\mathbf{R}^T\right)$$

To find the components $\mathbf{\Omega}'$ of the angular velocity vector on the principal axes, one just applies the formula given by Equation (7):

$$\mathbf{\Omega}' = \mathbf{R}^T\mathbf{C}^{-1}\left(\frac{d\mathbf{R}}{dt}\mathbf{R}^T\right) \tag{16}$$

This is the desired general relation, which we will now apply in practical situations.

Illustration: Euler angles

At this point, we have set up the general formalism to describe the rotational motion of rigid bodies. To go further, we have to specify generalized coordinates $\mathbf{q}$ that describe the motion of the body. One way of doing this is to use *Euler angles*.

Let us start by taking the reference orientation so that the principal axes coincide with the basis vectors $(\mathbf{e}_j)_{j=1,2,3} = (\hat{x}, \hat{y}, \hat{z})$. The Euler angles are based on the fact that any general rotation $\mathbf{R}$ can be written in terms of three angles (θ, ϕ, ψ) and so that $\mathbf{R}$ is the composition of three rotations

$$\mathbf{R}(\theta, \phi, \psi) = \mathbf{R}_z(\phi)\mathbf{R}_x(\theta)\mathbf{R}_z(\psi) \tag{17}$$

where $R_z(\psi)$ is a right-handed rotation of angle ψ around the z axis, $\mathbf{R}_x(\theta)$ is a right-handed rotation of angle θ about the x axis, and $\mathbf{R}_z(\phi)$ a right-handed rotation of angle ϕ about the z axis.

It can be shown that the Euler angles can be used to specify any orientation of a body. However, the angles for a given configuration are not unique: several triplets (θ, ϕ, ψ) can represent the same orientation. It is clear that when $\theta = 0$, all ϕ and ψ such that $\phi + \psi = \alpha$ for some given angle α will lead to the same orientation.

It is easy to compute $\mathbf{R}$ in terms of θ, ϕ and ψ. Indeed, for any angle β,

$$\mathbf{R}_z(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\mathbf{R}_x(\beta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix}$$

Computing $\mathbf{R}_z(\phi)\mathbf{R}_x(\theta)\mathbf{R}_z(\psi)$ is then straight forward, and we find

$$\mathbf{R}(\theta,\phi,\psi) = \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\sin\psi\cos\theta & -\cos\phi\sin\psi - \sin\phi\cos\psi\cos\theta & \sin\phi\sin\theta \\ \sin\phi\cos\psi + \cos\phi\sin\psi\cos\theta & \cos\phi\cos\psi\cos\theta - \sin\phi\sin\psi & -\cos\phi\sin\theta \\ \sin\psi\sin\theta & \cos\psi\sin\theta & \cos\theta \end{pmatrix}$$

Using patience or a computer that can handle symbolic manipulations, one can then evaluate

$$\frac{d\mathbf{R}}{dt}\mathbf{R}^T = \begin{pmatrix} 0 & -\frac{d\phi}{dt} - \frac{d\psi}{dt}\cos\theta & \sin\phi\frac{d\theta}{dt} - \frac{d\psi}{dt}\cos\phi\sin\theta \\ \frac{d\phi}{dt} + \frac{d\psi}{dt}\cos\theta & 0 & -\frac{d\psi}{dt}\sin\phi\sin\theta - \cos\phi\frac{d\theta}{dt} \\ \frac{d\psi}{dt}\cos\phi\sin\theta - \sin\phi\frac{d\theta}{dt} & \frac{d\psi}{dt}\sin\phi\sin\theta + \cos\phi\frac{d\theta}{dt} & 0 \end{pmatrix}$$

so that $\mathbf{C}^{-1}\left(\frac{d\mathbf{R}}{dt}\mathbf{R}^T\right)$ is the vector

$$\mathbf{C}^{-1}\left(\frac{d\mathbf{R}}{dt}\mathbf{R}^T\right) = \left(\frac{d\psi}{dt}\sin\phi\sin\theta + \cos\phi\frac{d\theta}{dt}, \sin\phi\frac{d\theta}{dt} - \frac{d\psi}{dt}\cos\phi\sin\theta, \frac{d\phi}{dt} + \frac{d\psi}{dt}\cos\theta\right)$$

and the components of the angular velocity vector along the principal axes are $\mathbf{\Omega}' = \mathbf{R}^T\mathbf{C}^{-1}\left(\frac{d\mathbf{R}}{dt}\mathbf{R}^T\right)$:

$$\mathbf{\Omega}' = \left(\frac{d\phi}{dt}\sin\psi\sin\theta + \frac{d\theta}{dt}\cos\psi, \frac{d\phi}{dt}\sin\theta\cos\psi - \frac{d\theta}{dt}\sin\psi, \frac{d\phi}{dt}\cos\theta + \frac{d\psi}{dt}\right) \tag{18}$$

There is a large degree of arbitrariness in the definition of the Euler angles. One can imagine several other constructions for the matrix $\mathbf{R}$, involving rotations around the y axis for example, and other angles. In addition, there is not much physical or mathematical justification for using these angles. The main reason why we chose to talk about them here is that it turns out they are convenient angles when describing the free rigid body (many would be) and the axisymmetric top.