Rotation of continuous bodies

August 31, 2024

1 Introduction

The rotation of continuous bodies tends to be the engineering/physic student's first introduction to a tensor. However, every text I've read makes what feels like a leap of faith in their derivation—the author knows what the answer is going to be, so he just pulls out the inertia tensor from thin air and uses the classic "well just check, it works" argument. It's the mathematical equivalent of solving a maze by getting 80% to the exit, then jumping to the exit and walking back to where you were and claiming you found the exit. True enough perhaps, but it's not a particularly satisfying pedagogical display. Why were you leading me to the exit if we were just going to jump to the end 80% through? What's more strange, it's only one extra line of mathematics that's needed to avoid making this leap of faith and just complete the derivation straight through. The main intention of this article is thus to add in that missing line to the derivation.

2 Rotation about the center of mass

Rotation about the center of mass is simpler to analyze than just an arbitrary axis and so we will start there. We begin with an inertial reference frame with a right-handed, orthonormal coordinate system. The x, y, and z axes' unit vectors respectively are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. The origin of this coordinate system we shall denote as N. Consider a body rotating with angular momentum $\boldsymbol{\omega}$ as seen in the inertial reference frame.

To find the angular momentum of a continuous body, we must start with the equation for the angular momentum of a particle. We can break the body into an infinite number of particles and sum each of the particle's angular momentum to find the total angular momentum of the body. To help facilitate this sum, we'll make a few definitions. First, let us attach point A to the body's center of mass. The position vector from the inertial coordinate system's origin to the body's center of mass is then $\mathbf{P}_{NA}$. We will denote the position vectors from the body's center of mass to each infinitesimal particle as $\mathbf{P}_{AB_i}$. The position vector from the origin to each infinitesimal mass is then

$$\mathbf{P}_{NB_i} = \mathbf{P}_{NA} + \mathbf{P}_{AB_i}. \tag{1}$$

The total angular momentum as observed in the inertial reference frame is therefore

$$\mathbf{L} = \sum_{i=1}^{\infty} \mathbf{P}_{AB_i} \times m_i \mathbf{V}_{NB_i} \tag{2}$$

where $\mathbf{V}_{NB_i}$ is the velocity of particle i as seen in the inertial reference frame, that is

$$\mathbf{V}_{NB_i} = \frac{d\mathbf{P}_{NB_i}}{dt}. \tag{3}$$

We can expand this velocity using Eq. (1).

$$\mathbf{V}_{NB_i} = \frac{d\mathbf{P}_{NA}}{dt} + \frac{d\mathbf{P}_{AB_i}}{dt} \tag{4}$$

and then we can use the transport theorem to transform the second velocity into the reference frame of the object as:

$$\mathbf{V}_{AB_i} = \frac{d\mathbf{P}_{AB_i}}{dt} = \frac{{}^{A}d\mathbf{P}_{AB_i}}{dt} + \boldsymbol{\omega} \times \mathbf{P}_{AB_i} = \mathbf{0} + \boldsymbol{\omega} \times \mathbf{P}_{AB_i} \tag{5}$$

where $\frac{{}^{A}d\mathbf{P}_{AB_i}}{dt}$ is the velocity of infinitesimal mass i as seen in the body attached frame, which is zero since in the body-fixed frame every point on the body is stationary (i.e. the body is not deforming). Substituting these results back into Eq. (2) yields

$$\mathbf{L} = \sum_{i=1}^{\infty} m_i \mathbf{P}_{AB_i} \times \frac{d\mathbf{P}_{NA}}{dt} + \sum_{i=1}^{\infty} \mathbf{P}_{AB_i} \times m_i \boldsymbol{\omega} \times \mathbf{P}_{AB_i}. \tag{6}$$

The first summation is zero which can be seen as follows. Notice the velocity is independent of the summing index in this term and so can be factored out of the sum:

$$\sum_{i=1}^{\infty} \left(m_i \mathbf{P}_{AB_i} \times \frac{d\mathbf{P}_{NA}}{dt} \right) = \sum_{i=1}^{\infty} (m_i \mathbf{P}_{AB_i}) \times \frac{d\mathbf{P}_{NA}}{dt} \tag{7}$$

Now, the summation on the right hand side is a sum of position vectors weighted by the mass at the position described by the vector. So, dividing by the total mass, we arrive at the formula for calculating the center of mass:

$$\frac{1}{M} \sum_{i=1}^{\infty} m_i \mathbf{P}_{AB_i}. \tag{8}$$

However, the sum being over the position vectors $\mathbf{P}_{AB_i}$ means that the result will be the center of mass with respect to the point A, which is already at the center of mass, so the entire sum is zero! Hence we may drop the first term from Eq. (6).

Next, we can rewrite the remaing sum's summand using the triple product expansion as

$$\mathbf{L} = \sum_{i=1}^{\infty} m_i \left[(\mathbf{P}_{AB_i} \cdot \mathbf{P}_{AB_i}) \boldsymbol{\omega} - (\mathbf{P}_{AB_i} \cdot \boldsymbol{\omega}) \mathbf{P}_{AB_i} \right]. \tag{9}$$

To continue any further, we must express $\mathbf{P}_{AB_i}$ and $\boldsymbol{\omega}$ in component form. The easiest choice is to use the body-fixed frame so that the position vectors are constant. Denoting the unit vectors of the x-, y-, and z-axes for the body-fixed frame as $\hat{\mathbf{x}}_A$, $\hat{\mathbf{y}}_A$, and $\hat{\mathbf{z}}_A$ respectively, we have

$$\mathbf{P}_{AB_i} = x_i\hat{\mathbf{x}}_A + y_i\hat{\mathbf{y}}_A + z_i\hat{\mathbf{z}}_A \tag{10}$$

$$\boldsymbol{\omega} = \omega_x\hat{\mathbf{x}}_A + \omega_y\hat{\mathbf{y}}_A + \omega_z\hat{\mathbf{z}}_A \tag{11}$$

The first term in the triple product is then

$$\begin{aligned}(\mathbf{P}_{AB_i}\cdot\mathbf{P}_{AB_i})\boldsymbol{\omega} &= \\ &(x_i^2\omega_x + y_i^2\omega_x + z_i^2\omega_x)\hat{\mathbf{x}}_A + (x_i^2\omega_y + y_i^2\omega_y + z_i^2\omega_y)\hat{\mathbf{y}}_A + \\ &(x_i^2\omega_z + y_i^2\omega_z + z_i^2\omega_z)\hat{\mathbf{z}}_A\end{aligned} \tag{12}$$

while the second term is

$$\begin{aligned}-(\mathbf{P}_{AB_i}\cdot\boldsymbol{\omega})\mathbf{P}_{AB_i} &= \\ &(x_i^2\omega_x + y_ix_i\omega_y + z_ix_i\omega_z)\hat{\mathbf{x}}_A + (x_iy_i\omega_x + y_i^2\omega_y + z_iy_i\omega_z)\hat{\mathbf{y}}_A + \\ &(x_iz_i\omega_x + y_iz_i\omega_y + z_i^2\omega_z)\hat{\mathbf{z}}_A\end{aligned} \tag{13}$$

thus the full triple product is

$$\begin{aligned}(\mathbf{P}_{AB_i}\cdot\mathbf{P}_{AB_i})\boldsymbol{\omega} - (\mathbf{P}_{AB_i}\cdot\boldsymbol{\omega})\mathbf{P}_{AB_i} &= \\ &((y_i^2 + z_i^2)\omega_x - y_ix_i\omega_y - z_ix_i\omega_z)\hat{\mathbf{x}}_A \\ &+ ((x_i^2 + z_i^2)\omega_y - x_iy_i\omega_x - z_iy_i\omega_z)\hat{\mathbf{y}}_A \\ &+ ((x_i^2 + y_i^2)\omega_z - x_iz_i\omega_x - y_iz_i\omega_y)\hat{\mathbf{z}}_A.\end{aligned} \tag{14}$$

Now comes the most crucial step, which is omitted in every derivation I've seen. We use the fact that $\omega_x = (\boldsymbol{\omega}\cdot\hat{\mathbf{x}}_A)$, and similarly for the y- and z-components, to reinsert $\boldsymbol{\omega}$ in its vector form viz.

$$\begin{aligned}(\mathbf{P}_{AB_i}\cdot\mathbf{P}_{AB_i})\boldsymbol{\omega} - (\mathbf{P}_{AB_i}\cdot\boldsymbol{\omega})\mathbf{P}_{AB_i} &= \\ &(y_i^2 + z_i^2)(\boldsymbol{\omega}\cdot\hat{\mathbf{x}}_A)\hat{\mathbf{x}}_A - y_ix_i(\boldsymbol{\omega}\cdot\hat{\mathbf{y}}_A)\hat{\mathbf{x}}_A - z_ix_i(\boldsymbol{\omega}\cdot\hat{\mathbf{z}}_A)\hat{\mathbf{x}}_A \\ &+ (x_i^2 + z_i^2)(\boldsymbol{\omega}\cdot\hat{\mathbf{y}}_A)\hat{\mathbf{y}}_A - x_iy_i(\boldsymbol{\omega}\cdot\hat{\mathbf{x}}_A)\hat{\mathbf{y}}_A - z_iy_i(\boldsymbol{\omega}\cdot\hat{\mathbf{z}}_A)\hat{\mathbf{z}}_A \\ &+ (x_i^2 + y_i^2)(\boldsymbol{\omega}\cdot\hat{\mathbf{z}}_A)\hat{\mathbf{z}}_A - x_iz_i(\boldsymbol{\omega}\cdot\hat{\mathbf{x}}_A)\hat{\mathbf{z}}_A - y_iz_i(\boldsymbol{\omega}\cdot\hat{\mathbf{y}}_A)\hat{\mathbf{z}}_A\end{aligned} \tag{15}$$

where I have distributed the unit vectors out in anticipation of the next step. Notice that all nine terms have a dot product of $\boldsymbol{\omega}$ on the left, thus we can factor out $\boldsymbol{\omega}$ to arrive at

$$\begin{aligned}(\mathbf{P}_{AB_i}\cdot\mathbf{P}_{AB_i})\boldsymbol{\omega} - (\mathbf{P}_{AB_i}\cdot\boldsymbol{\omega})\mathbf{P}_{AB_i} &= \\ &\boldsymbol{\omega}\cdot[(y_i^2 + z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - y_ix_i\hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - z_ix_i\hat{\mathbf{z}}_A\hat{\mathbf{x}}_A \\ &+ (x_i^2 + z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - x_iy_i\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - z_iy_i\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A \\ &+ (x_i^2 + y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - x_iz_i\hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - y_iz_i\hat{\mathbf{y}}_A\hat{\mathbf{z}}_A]\end{aligned} \tag{16}$$

The terms $\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A$, $\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A$ etc. are known as dyads. Order matters with dyads, so $\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A \neq \hat{\mathbf{y}}_A\hat{\mathbf{x}}_A$. The side we take a dot product from matters also. For instance, if we want to dot $\hat{\mathbf{x}}_A$ into $\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A$ from the left we have $\hat{\mathbf{x}}_A\cdot\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A = (\hat{\mathbf{x}}_A\cdot\hat{\mathbf{x}}_A)\hat{\mathbf{y}}_A = \hat{\mathbf{y}}_A$ while dotting from the right gives us $\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A\cdot\hat{\mathbf{x}}_A = \hat{\mathbf{x}}_A(\hat{\mathbf{y}}_A\cdot\hat{\mathbf{x}}_A) = \mathbf{0}$.

Now, inserting Eq. (17) into Eq. (9) allows us to arrive at the inertia tensor:

$$\begin{aligned}\mathbf{L} = \sum_{i=1}^{\infty}\boldsymbol{\omega}\cdot[&m_i(y_i^2+z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - m_iy_ix_i\hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - m_iz_ix_i\hat{\mathbf{z}}_A\hat{\mathbf{x}}_A\\ &+ m_i(x_i^2+z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - m_ix_iy_i\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - m_iz_iy_i\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A\\ &+ m_i(x_i^2+y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - m_ix_iz_i\hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - m_iy_iz_i\hat{\mathbf{y}}_A\hat{\mathbf{z}}_A]\end{aligned} \tag{17}$$

and since $\boldsymbol{\omega}$ is constant for each term of the sum, it may be factored out:

$$\begin{aligned}\mathbf{L} = \boldsymbol{\omega}\cdot\sum_{i=1}^{\infty}[&m_i(y_i^2+z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - m_iy_ix_i\hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - m_iz_ix_i\hat{\mathbf{z}}_A\hat{\mathbf{x}}_A\\ &+ m_i(x_i^2+z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - m_ix_iy_i\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - m_iz_iy_i\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A\\ &+ m_i(x_i^2+y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - m_ix_iz_i\hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - m_iy_iz_i\hat{\mathbf{y}}_A\hat{\mathbf{z}}_A] = \boldsymbol{\omega}\cdot\mathbf{I}\end{aligned} \tag{18}$$

and in the final equality we have identified the moment of inertia dyad $\mathbf{I}$ as

$$\begin{aligned}\mathbf{I} = \sum_{i=1}^{\infty}[&m_i(y_i^2+z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - m_iy_ix_i\hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - m_iz_ix_i\hat{\mathbf{z}}_A\hat{\mathbf{x}}_A\\ &+ m_i(x_i^2+z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - m_ix_iy_i\hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - m_iz_iy_i\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A\\ &+ m_i(x_i^2+y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - m_ix_iz_i\hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - m_iy_iz_i\hat{\mathbf{y}}_A\hat{\mathbf{z}}_A]\end{aligned} \tag{19}$$

Now, the equation for angular moment here is backwards from what is usually shown in textbooks which is $\mathbf{L} = \mathbf{I}\cdot\boldsymbol{\omega}$. The question naturally arises then, what is the difference and why are both valid if left and right dot products are not generally equal for a dyad? The difference is that the moment of inertia as I have defined it is the transpose of the moment of inertia dyad as it is usually defined. The upshot is that the moment of inertia dyad is symmetric so dotting from the left and right yields the same result (this all will be proved shortly). I did this because it is pedagoically easier to arrive at this result. If we wanted to factor out omega from the right hand side we could have done so as well, but it requires a bit more work. Instead, at Eq. (15) we would insert $\hat{\mathbf{x}}_A\cdot\boldsymbol{\omega}$ for ω_x rather than $\boldsymbol{\omega}\cdot\hat{\mathbf{x}}_A$ and similarly for the y- and z-components. This would lead to

$$\begin{aligned}(\mathbf{P}_{AB_i}\cdot\mathbf{P}_{AB_i})\boldsymbol{\omega} - (\mathbf{P}_{AB_i}\cdot\boldsymbol{\omega})\mathbf{P}_{AB_i} =&\\ (y_i^2+z_i^2)(\hat{\mathbf{x}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{x}}_A &- y_ix_i(\hat{\mathbf{y}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{x}}_A - z_ix_i(\hat{\mathbf{z}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{x}}_A\\ + (x_i^2+z_i^2)(\hat{\mathbf{y}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{y}}_A &- x_iy_i(\hat{\mathbf{x}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{y}}_A - z_iy_i(\hat{\mathbf{z}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{z}}_A\\ + (x_i^2+y_i^2)(\hat{\mathbf{z}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{z}}_A &- x_iz_i(\hat{\mathbf{x}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{z}}_A - y_iz_i(\hat{\mathbf{y}}_A\cdot\boldsymbol{\omega})\hat{\mathbf{z}}_A.\end{aligned} \tag{20}$$

To factor out $\boldsymbol{\omega}$ from here, we must recognize that since $(\hat{\mathbf{x}}_A \cdot \boldsymbol{\omega})$ etc. are all scalars, $(\hat{\mathbf{x}}_A \cdot \boldsymbol{\omega})\hat{\mathbf{x}}_A$ is the same as $\hat{\mathbf{x}}_A(\hat{\mathbf{x}}_A \cdot \boldsymbol{\omega})$ and similarly for the other 8 terms. Then Eq. (20) becomes

$$\begin{aligned}(\mathbf{P}_{AB_i} \cdot \mathbf{P}_{AB_i})\boldsymbol{\omega} - (\mathbf{P}_{AB_i} \cdot \boldsymbol{\omega})\mathbf{P}_{AB_i} = &\\ [(y_i^2 + z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - y_i x_i \hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - z_i x_i \hat{\mathbf{z}}_A\hat{\mathbf{x}}_A &\\ + (x_i^2 + z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - x_i y_i \hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - z_i y_i \hat{\mathbf{z}}_A\hat{\mathbf{z}}_A &\\ + (x_i^2 + y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - x_i z_i \hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - y_i z_i \hat{\mathbf{y}}_A\hat{\mathbf{z}}_A] \cdot \boldsymbol{\omega}. & \quad (21)\end{aligned}$$

This can be reinserted into Eq. (9) to arrive at

$$\mathbf{L} = \boldsymbol{\Psi} \cdot \boldsymbol{\omega} \quad (22)$$

where $\boldsymbol{\Psi}$ is defined as

$$\begin{aligned}\boldsymbol{\Psi} = \sum_{i=1}^{\infty} m_i[(y_i^2 + z_i^2)\hat{\mathbf{x}}_A\hat{\mathbf{x}}_A - y_i x_i \hat{\mathbf{y}}_A\hat{\mathbf{x}}_A - z_i x_i \hat{\mathbf{z}}_A\hat{\mathbf{x}}_A &\\ + (x_i^2 + z_i^2)\hat{\mathbf{y}}_A\hat{\mathbf{y}}_A - x_i y_i \hat{\mathbf{x}}_A\hat{\mathbf{y}}_A - z_i y_i \hat{\mathbf{z}}_A\hat{\mathbf{z}}_A &\\ + (x_i^2 + y_i^2)\hat{\mathbf{z}}_A\hat{\mathbf{z}}_A - x_i z_i \hat{\mathbf{x}}_A\hat{\mathbf{z}}_A - y_i z_i \hat{\mathbf{y}}_A\hat{\mathbf{z}}_A]. & \quad (23)\end{aligned}$$

Finally, comparing this equation to that of Eq. (19) term by term we see that $\boldsymbol{\Psi}$ is equal to that of $\mathbf{I}$ leading to the equality $\Psi = \mathbf{I}$ and hence $\mathbf{L} = \boldsymbol{\omega} \cdot \mathbf{I} = \mathbf{I} \cdot \boldsymbol{\omega}$ so for the inertia dyad, the left and right dot products are equivalent.