2.003J Spring 2011: Dynamics and Control I Massachusetts Institute of Technology Department of Mechanical Engineering	**On Notation** **Jason Ku** **Feb. 08, 2011**

On Notation

Much of 2.003j involves defining then manipulating points, frames of reference, and vectors. Our ability to precisely communicate how these objects interact with each other without ambiguity rests on our choice of notation. In this course, you will encounter subtly different notation in lectures, in handouts, and in your book. We will try to be consistent in our notation in order to be as clear as possible.

Points

Points will appear in bold, scripted letters. This notation should be consistent in lecture, handouts, and in the book.

Example notation of a point $\Longrightarrow (P, O, A, o, q)$.

Frames of Reference

Frames of reference are very important in 2.003j. For example, a velocity is meaningless unless it can be compared to some frame of reference. Reference frames are defined and labeled with respect to an origin point. For example, the reference frame $\hat{\mathrm{O}}$ by definition has its origin at point O. Note that we can define two different reference frames $\hat{\mathrm{O}} \neq \hat{\mathrm{A}}$ that rotate with respect to one another but share the same origin, $O \equiv A$. Reference frames have the same style as points but with a hat on top to remind us that we are talking about a frame of reference. The letters I, J, and K will be reserved for unit coordinate vectors so, for clarity, should not be used to label reference frames.

Example notation of a reference frame $\Longrightarrow \left(\hat{\mathrm{O}}, \hat{\mathrm{o}}, \hat{\mathrm{A}}, \hat{\mathrm{B}}\right)$

Unit Vectors

Unit vectors $\hat{\mathbf{i}}\hat{\mathbf{j}}\hat{\mathbf{k}}$ for our reference frames will always form a right handed triad with $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$ and will appear bold and unscripted with a hat on top. To denote that a coordinate system unit vector is attached to a specific reference frame, say frame $\hat{\mathrm{A}}$, we will use a subscript on the right of the vector, or $\hat{\mathbf{i}}_{\mathrm{A}}$. We will remove the hat from the subscript as it is clear from the context that it is a reference frame. In addition, we will sometimes use capital letters to denote the unit coordinate vectors in the inertial ground reference frame, $\hat{\mathbf{i}}_{\mathrm{G}} = \hat{\mathbf{I}}$ for convenience. In lecture, unit coordinate vectors will not be bold, but the meaning should remain clear.

Example notation of a unit coordinate vector $\Longrightarrow \left(\hat{\mathbf{i}}, \hat{\mathbf{J}}, \hat{\mathbf{j}}_{\mathrm{O}}, \hat{\mathbf{k}}_{\mathrm{A}}\right)$ or $\left(\hat{\mathrm{i}}, \hat{\mathrm{J}}, \hat{\mathrm{j}}_{\mathrm{O}}, \hat{\mathrm{k}}_{\mathrm{A}}\right)$ in lecture

Vectors

Vectors describe the position, velocity, or acceleration of points in your system and will appear bold or, in lecture, with an arrow across the top. The position, velocity, and acceleration of a point must be defined with respect to some frame of reference. Without the context of a reference frame, these vectors would be ambiguous. For example, the velocity vector $\boldsymbol{v}$ of point p with respect to frame $\hat{\mathrm{A}}$ will be denoted as ${}^{\mathrm{A}}\boldsymbol{v}_p$. Again we will drop the hat on the reference frame in the superscript as it is clear from the context that it is a reference frame. Note that the book does not typically use this precise notation, but we will ask you to use it for the sake of clarity and precision.

$$\textbf{Example notation of a movement vector} \Longrightarrow \left({}^{\mathrm{O}}\boldsymbol{r}_p, {}^{\mathrm{A}}\boldsymbol{a}_o, {}^{\mathrm{B}}\boldsymbol{v}_q\right) \text{ or } \left({}^{\mathrm{O}}\vec{r}_p, {}^{\mathrm{A}}\vec{a}_o, {}^{\mathrm{B}}\vec{v}_q\right) \text{ in lecture}$$

Rotations

Rotations refer to the rotation of a reference frame with respect to another reference frame, as the rotation of a one dimensional point has no useful meaning. While a rotation is not a vector, angular velocity and angular acceleration are. For example, the angular velocity $\boldsymbol{\omega}$ of frame $\hat{\mathrm{A}}$ with respect to frame $\hat{\mathrm{B}}$ will be denoted as ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}}$. Again we will drop the hat on the reference frames for convenience.

$$\textbf{Example notation of a movement vector} \Longrightarrow \left({}^{\mathrm{G}}\boldsymbol{\omega}_{\mathrm{A}}, {}^{\mathrm{G}}\boldsymbol{\Omega}_{\mathrm{B}}, {}^{\mathrm{A}}\boldsymbol{\phi}_{\mathrm{B}}\right) \text{ or } \left({}^{\mathrm{G}}\vec{\omega}_{\mathrm{A}}, {}^{\mathrm{G}}\vec{\Omega}_{\mathrm{B}}, {}^{\mathrm{A}}\vec{\phi}_{\mathrm{B}}\right) \text{ in lecture}$$

Time Derivatives

Time derivatives must always be taken with respect to a reference frame. The time derivative $\frac{d}{dt}$ with respect to frame $\hat{\mathrm{A}}$ will be denoted as $\frac{{}^{\mathrm{A}}d}{dt}$. Note that the book uses the notations $\frac{d}{dt}$ and $\frac{d}{dt}()_{rel}$ which can be both confusing and ambiguous. The "dot" operator $\dot{\boldsymbol{r}}$ will denote the time derivative with respect to the frame that $\boldsymbol{r}$ is defined in, with $\frac{{}^{\mathrm{A}}d}{dt}({}^{\mathrm{A}}\boldsymbol{r}_p) = {}^{\mathrm{A}}\dot{\boldsymbol{r}}_p$.

$$\textbf{Example notation of time derivatives} \Longrightarrow \left(\frac{{}^{\mathrm{G}}d}{dt}\left({}^{\mathrm{A}}\boldsymbol{r}_p\right), {}^{\mathrm{A}}\dot{\boldsymbol{\omega}}_B\right) \text{ or } \left(\frac{{}^{\mathrm{G}}d}{dt}\left({}^{\mathrm{A}}\vec{r}_p\right), {}^{\mathrm{A}}\dot{\vec{\omega}}_{\mathrm{B}}\right) \text{ in lecture}$$

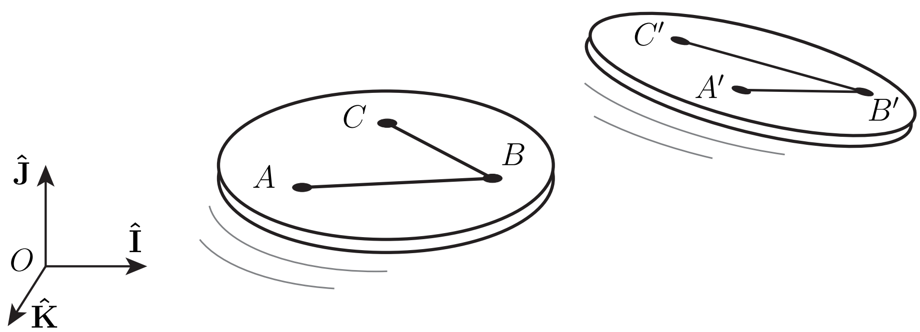
Relating Reference Frames

This notation may seem bulky and cumbersome at first but should make dealing with multiple reference frames clearer and more precise. For example, define an inertial ground reference frame $\hat{\mathrm{O}}$ at origin O and reference frame $\hat{\mathrm{A}}$ with origin A with angular velocity ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}}$ with respect to frame $\hat{\mathrm{O}}$. Then for any point p:

$$\frac{{}^{\mathrm{O}}d}{dt}\left({}^{\mathrm{O}}\boldsymbol{r}_p\right) = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_p = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_A + \frac{{}^{\mathrm{O}}d}{dt}\left({}^{\mathrm{A}}\boldsymbol{r}_p\right) = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\dot{\boldsymbol{r}}_p + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_p$$

Derivation of Velocity and Acceleration Equations Using Intermediate Frames

Why should we use intermediate reference frames when we approach kinematic problems? Thinking in terms of intermediate reference frames can often simplify our calculations, while attaching different intermediate reference frames to different rigid bodies of our system can provide some physical intuition about their relative motions.



For example, a rigid body like a frisbee might fly through the air while translating and rotating in a very complicated motion. However, given three points A, B, and C that are rigidly attached to the frisbee, the distances $\overline{AB}$, $\overline{BC}$, and $\overline{CA}$, as well as the three dimensional angles $\angle ABC$, $\angle BCA$, and $\angle CAB$ all must be constant in time, regardless of our frame of reference. This is a powerful property and using intermediate reference frames can help us exploit this property.

In order to work with reference frames, we must mathematically define how the derivative operator applies with respect to different frames of reference. First, we will derive the derivative of a vector with respect to the frame in which it is defined. This result is familiar and is consistent with the traditional definition of the derivative of a vector. Then we will derive the derivative of a vector with respect to an intermediate reference frame, first without rotation, then finally allowing rotation of the intermediate frame.

Derivation of Velocity and Acceleration Equations Using Intermediate Frames

Jason Ku, Feb. 12, 2011

1. Derivative of a Vector in Its Own Reference Frame $\frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{O}}\boldsymbol{r}_P)$

Define fixed reference frame $\hat{\mathrm{O}} = (O, \hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$ and arbitrary point P. Define the vector ${}^{\mathrm{O}}\boldsymbol{r}_P = x\hat{\mathbf{i}}_O + y\hat{\mathbf{j}}_O + z\hat{\mathbf{k}}_O$, where the values (x, y, z) are the scalar projections of vector ${}^{\mathrm{O}}\boldsymbol{r}_P$ on the unit coordinate vectors of $\hat{\mathrm{O}}$. For example $x = {}^{\mathrm{O}}\boldsymbol{r}_P \cdot \hat{\mathbf{i}}_O$. Let us take the derivative of ${}^{\mathrm{O}}\boldsymbol{r}_P$ with respect to the reference frame $\hat{\mathrm{O}}$ in which it is defined by using the chain rule:

$$
\begin{aligned}
\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{O}}\boldsymbol{r}_P &= \frac{{}^{\mathrm{O}}d}{dt}(x\hat{\mathbf{i}}_O + y\hat{\mathbf{j}}_O + z\hat{\mathbf{k}}_O) \\
&= \frac{dx}{dt}\hat{\mathbf{i}}_O + \frac{dy}{dt}\hat{\mathbf{j}}_O + \frac{dz}{dt}\hat{\mathbf{k}}_O + x\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_O} + y\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_O} + z\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_O} \\
&= {}^{\mathrm{O}}\dot{\boldsymbol{r}}_P
\end{aligned}
$$

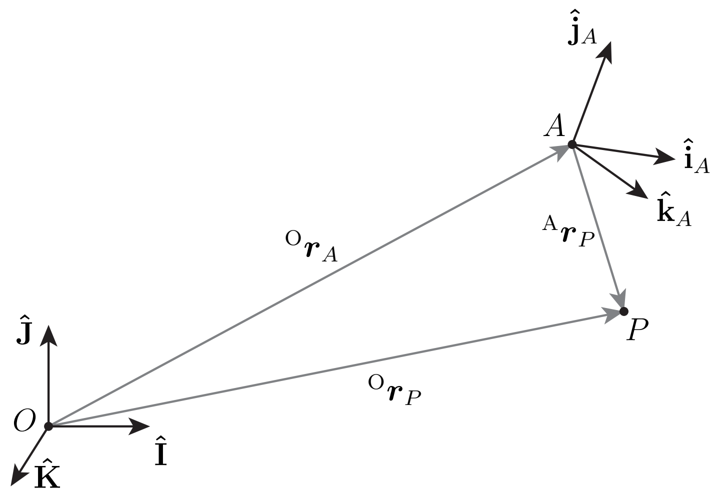
Here we exploit the usefulness of reference frames. The unit vectors of frame $\hat{\mathrm{O}}$ by definition do not change in time with respect to the $\hat{\mathrm{O}}$ reference frame. Also note that derivatives of scalar values do not depend on a frame of reference. This is the traditional definition of the derivative of a vector, and this is what we mean when we use the "dot" operator.

2. Derivative of a Vector in an Intermediate Frame Without Rotation $\frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{A}}\boldsymbol{r}_P)$

Let us define two reference frames (a fixed frame and an intermediate frame) which only translate with respect to each other. Define fixed reference frame $\hat{\mathrm{O}} = (O, \hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$, intermediate frame $\hat{\mathrm{A}} = (A, \hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$, and arbitrary point P. For pure translation, the origin of frame $\hat{\mathrm{A}}$ may change in time with respect to frame $\hat{\mathrm{O}}$, but the orientation of the unit vectors $(\hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$ should not change in time with respect to frame $\hat{\mathrm{O}}$, or mathematically, $\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A = \frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A = \frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A = 0$. Again define the vector ${}^{\mathrm{A}}\boldsymbol{r}_P = x\hat{\mathbf{i}}_A + y\hat{\mathbf{j}}_A + z\hat{\mathbf{k}}_A$. Let us take the derivative of ${}^{\mathrm{A}}\boldsymbol{r}_P$ with respect to the reference frame $\hat{\mathrm{O}}$, again using the chain rule:

$$
\begin{aligned}
\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P &= \frac{{}^{\mathrm{O}}d}{dt}(x\hat{\mathbf{i}}_A + y\hat{\mathbf{j}}_A + z\hat{\mathbf{k}}_A) \\
&= \frac{dx}{dt}\hat{\mathbf{i}}_A + \frac{dy}{dt}\hat{\mathbf{j}}_A + \frac{dz}{dt}\hat{\mathbf{k}}_A + x\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A} + y\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A} + z\cancelto{0}{\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A} \\
&= {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P
\end{aligned}
$$

Interestingly, whether we take the derivative of ${}^{\mathrm{A}}\boldsymbol{r}_P$ with respect to frame $\hat{\mathrm{O}}$ or frame $\hat{\mathrm{A}}$, our result is the same under pure translation. Here we exploit the fact that the derivatives of the unit vectors in a purely translating frame must be zero. Note also that if ${}^{\mathrm{A}}\boldsymbol{r}_P$ is fixed in frame $\hat{\mathrm{A}}$, then by definition, ${}^{\mathrm{A}}\dot{\boldsymbol{r}}_P = 0$.



3. Derivative of a Vector in an Intermediate Frame With Rotation $\frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{A}}\boldsymbol{r}_P)$

Rotations are more complicated and a bit less intuitive. However, we can use the same techniques to derive a derivative. Let us define two reference frames (a fixed frame and an intermediate frame) which can both translate and rotate with respect to each other. Define fixed reference frame $\hat{\mathrm{O}} = (O, \hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$, intermediate frame $\hat{\mathrm{A}} = (A, \hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$, and arbitrary point P. Again define the vector ${}^{\mathrm{A}}\boldsymbol{r}_P = x\hat{\mathbf{i}}_A + y\hat{\mathbf{j}}_A + z\hat{\mathbf{k}}_A$. Now, let us take the derivative of ${}^{\mathrm{A}}\boldsymbol{r}_P$ with respect to the reference frame $\hat{\mathrm{O}}$, again using the chain rule:

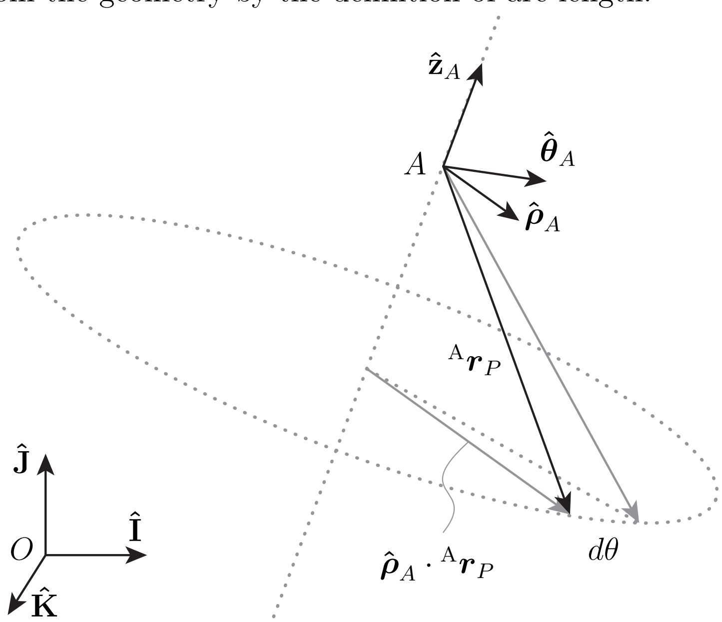
$$\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P = \frac{{}^{\mathrm{O}}d}{dt}(x\hat{\mathbf{i}}_A + y\hat{\mathbf{j}}_A + z\hat{\mathbf{k}}_A)$$

$$= \frac{dx}{dt}\hat{\mathbf{i}}_A + \frac{dy}{dt}\hat{\mathbf{j}}_A + \frac{dz}{dt}\hat{\mathbf{k}}_A + x\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A + y\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A + z\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A$$

$$= {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + \left(x\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A + y\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A + z\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A\right)$$

Unlike before, frames $\hat{\mathrm{O}}$ and $\hat{\mathrm{A}}$ rotate with respect to one another, so the direction of the unit vectors $(\hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$ may change in time with respect to frame $\hat{\mathrm{O}}$. Let's look closer at the concept of rotation. We know that frame $\hat{\mathrm{A}}$ rotates with respect to fixed frame $\hat{\mathrm{O}}$ with some angular velocity about an axis of rotation passing through A. Assume for the moment that ${}^{\mathrm{A}}\boldsymbol{r}_P$ is fixed in frame $\hat{\mathrm{A}}$. Then by definition, ${}^{\mathrm{A}}\dot{\boldsymbol{r}}_P = 0$. The only way the direction of this vector could change with respect to frame $\hat{\mathrm{O}}$ would be because of an instantaneous rotation.

For convenience, let us redefine coordinate system $\hat{\text{A}}$ using a different set of unit vectors reoriented with respect to both the axis of rotation, which in general could point in any arbitrary direction, and the vector ${}^{\text{A}}\boldsymbol{r}_P$. Define $\hat{\mathbf{z}}_A$ to be parallel to the axis of rotation, $\hat{\boldsymbol{\rho}}_A$ to be the radial coordinate, and $\hat{\boldsymbol{\theta}}_A$ to be the tangential coordinate, such that ${}^{\text{A}}\boldsymbol{r}_P \cdot \hat{\boldsymbol{\theta}}_A = 0$. It is essential to note that we have not defined a new reference frame, but instead re-expressed $\hat{\text{A}}$ using different unit vectors. This is a right handed system with $\hat{\boldsymbol{\theta}}_A \times \hat{\mathbf{z}}_A = \hat{\boldsymbol{\rho}}_A$. After a small change in time dt, vector ${}^{\text{A}}\boldsymbol{r}_P$ rotates by angle $d\theta$. The following relation holds from the geometry by the definition of arc length:



$$ {}^{O}d({}^{\text{A}}\boldsymbol{r}_P) = d\theta(\hat{\boldsymbol{\rho}}_A \cdot {}^{\text{A}}\boldsymbol{r}_P)\hat{\boldsymbol{\theta}}_A $$

The infinitesimal change in the unit vector ${}^{\text{A}}\boldsymbol{r}_P$ should be equal to the infinitesimal change in angle $d\theta$ times the radial component of ${}^{\text{A}}\boldsymbol{r}_P$ and should be pointed in the $\hat{\boldsymbol{\theta}}_A$ direction for positive $d\theta$. Since $\hat{\boldsymbol{\theta}}_A \times \hat{\mathbf{z}}_A = \hat{\boldsymbol{\rho}}_A$, we can write:

$$ {}^{O}d({}^{\text{A}}\boldsymbol{r}_P) = d\theta((\hat{\boldsymbol{\theta}}_A \times \hat{\mathbf{z}}_A) \cdot {}^{\text{A}}\boldsymbol{r}_P)\hat{\boldsymbol{\theta}}_A $$

Exploiting the vector identity that $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ we can write:

$$ {}^{O}d({}^{\text{A}}\boldsymbol{r}_P) = d\theta((\hat{\mathbf{z}}_A \times {}^{\text{A}}\boldsymbol{r}_P) \cdot \hat{\boldsymbol{\theta}}_A)\hat{\boldsymbol{\theta}}_A $$

Also, defined by our choice of unit coordinates, ${}^{\text{A}}\boldsymbol{r}_P \cdot \hat{\boldsymbol{\theta}}_A = 0$, so $((\hat{\mathbf{z}}_A \times {}^{\text{A}}\boldsymbol{r}_P) \cdot \hat{\boldsymbol{\theta}}_A)\hat{\boldsymbol{\theta}}_A = (\hat{\mathbf{z}}_A \times {}^{\text{A}}\boldsymbol{r}_P)$. Dividing both sides by dt yields:

$$ \begin{aligned} \frac{{}^{\text{O}}d}{dt}{}^{\text{A}}\boldsymbol{r}_P &= \frac{d\theta}{dt}\hat{\mathbf{z}}_A \times {}^{\text{A}}\boldsymbol{r}_P \\ &= {}^{\text{O}}\boldsymbol{\omega}_{\text{A}} \times {}^{\text{A}}\boldsymbol{r}_P \end{aligned} $$

Here we define ${}^{\text{O}}\boldsymbol{\omega}_{\text{A}}$ to be equal to the instantaneous rate of change of the rotation angle in the direction parallel to the axis of rotation $\frac{d\theta}{dt}\hat{\mathbf{z}}$ given by the right hand rule.

Our only assumption in the previous analysis was for ${}^{\mathrm{A}}\boldsymbol{r}_P$ to be fixed in reference frame $\hat{\mathrm{A}}$, so the term ${}^{\mathrm{A}}\dot{\boldsymbol{r}}_P = 0$. Thus, when ${}^{\mathrm{A}}\boldsymbol{r}_P$ is fixed in frame $\hat{\mathrm{A}}$, we can conclude that:

$$\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P = \cancelto{0}{{}^{\mathrm{A}}\dot{\boldsymbol{r}}_P} + \left(x\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A + y\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A + z\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A \right)$$
$$= x\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{i}}_A + y\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{j}}_A + z\frac{{}^{\mathrm{O}}d}{dt}\hat{\mathbf{k}}_A = {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P$$

On the other hand, if ${}^{\mathrm{A}}\boldsymbol{r}_P$ is not fixed in frame $\hat{\mathrm{A}}$, ${}^{\mathrm{A}}\dot{\boldsymbol{r}}_P \neq 0$. Thus the general formula for taking the time derivative of any vector ${}^{\mathrm{A}}\boldsymbol{r}_P$ with respect to a different frame $\hat{\mathrm{O}}$ is:

$$\boxed{\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P = {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P}$$

Note that this formula is completely general and valid when taking the time derivative of any vector with respect to any other reference frame. For example, if we take our time derivative with respect to the same reference frame that the vector is defined in (as in Section 1), ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{O}} = \mathbf{0}$ by definition, and we yield the usual result:

$$\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{O}}\boldsymbol{r}_P = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_P + \cancelto{0}{{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{O}}} \times {}^{\mathrm{O}}\boldsymbol{r}_P = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_P$$

4. Derivation of Velocity and Acceleration Using an Intermediate Frame

Quite often, we must find the velocity and acceleration of a vector, but be given values defined with respect to intermediate frames. Let us define two reference frames (a fixed frame and an intermediate frame) which can both translate and rotate with respect to each other. Define fixed reference frame $\hat{\mathrm{O}} = (O, \hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$, intermediate frame $\hat{\mathrm{A}} = (A, \hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$, and arbitrary point P. Now that we have formula for the time derivative of a vector with respect to any frame, we can proceed directly:

$${}^{\mathrm{O}}\boldsymbol{v}_P = \frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{O}}\boldsymbol{r}_P = \frac{{}^{\mathrm{O}}d}{dt}\left({}^{\mathrm{O}}\boldsymbol{r}_A + {}^{\mathrm{A}}\boldsymbol{r}_P\right) = {}^{\mathrm{O}}\dot{\boldsymbol{r}}_A + \frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P = \boxed{{}^{\mathrm{O}}\dot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P}$$

For acceleration, we just take an additional time derivative with respect to frame $\hat{\mathrm{O}}$:

$${}^{\mathrm{O}}\boldsymbol{a}_P = \frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{O}}\boldsymbol{v}_P = \frac{{}^{\mathrm{O}}d}{dt}\left({}^{\mathrm{O}}\dot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P\right)$$
$$= {}^{\mathrm{O}}\ddot{\boldsymbol{r}}_A + \frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + \left(\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}}\right) \times {}^{\mathrm{A}}\boldsymbol{r}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times \left(\frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P\right)$$
$$= {}^{\mathrm{O}}\ddot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\ddot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times \left({}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P\right)$$
$$= \boxed{{}^{\mathrm{O}}\ddot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\ddot{\boldsymbol{r}}_P + 2\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times \left({}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P\right)}$$

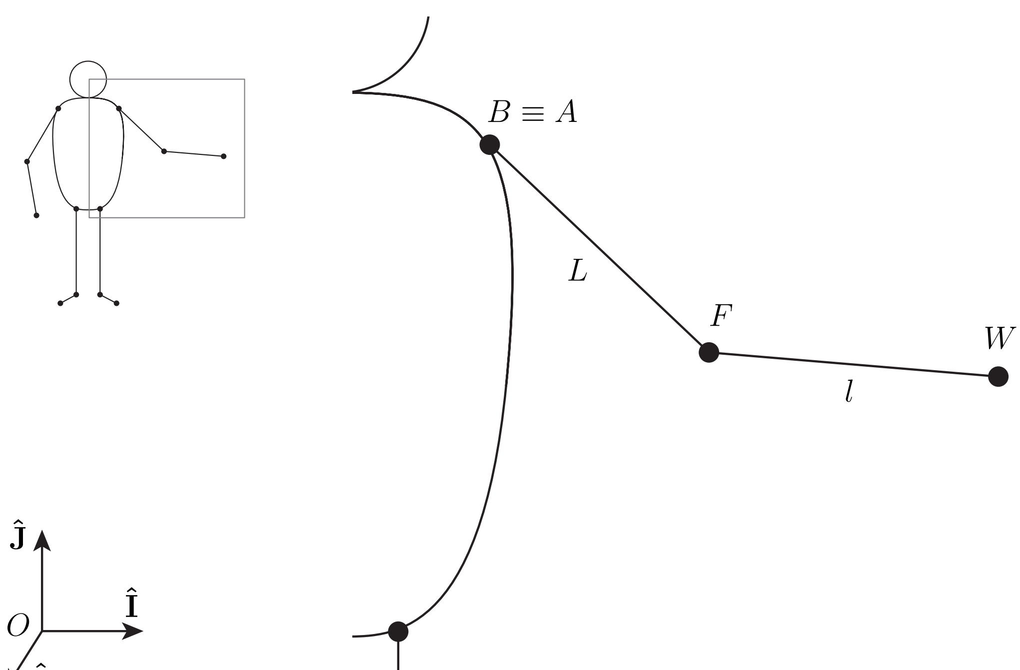
5. Physical Meaning of Terms in the Acceleration Equation

$$ {}^{\mathrm{O}}\boldsymbol{a}_P = {}^{\mathrm{O}}\ddot{\boldsymbol{r}}_A + {}^{\mathrm{A}}\ddot{\boldsymbol{r}}_P + 2\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times \left({}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P\right) $$

Each term in this equation has a physical meaning. The first two terms are fairly intuitive: ${}^{\mathrm{O}}\ddot{\boldsymbol{r}}_A$ is the acceleration of point A with respect to frame $\hat{\mathrm{O}}$, and ${}^{\mathrm{A}}\ddot{\boldsymbol{r}}_P$ is the acceleration of point P with respect to frame $\hat{\mathrm{A}}$. The remaining three terms result from the fact that frame $\hat{\mathrm{A}}$ and frame $\hat{\mathrm{O}}$ rotate with respect to one another. The term $2\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P$ results if point P moves relative to the rotating frame $\hat{\mathrm{A}}$. For example, you experience this acceleration when you walk around while riding on a spinning carousel. This is called the Coriolis acceleration and, unlike the other terms, has a coefficient of 2. The term ${}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P$ results if ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}}$ changes in time. For example, you experience this acceleration when you rotate a gyroscope, changing the orientation of its angular velocity. This is called the Eulerian acceleration. The last term ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times \left({}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P\right)$ is probably the term most familiar to you and is necessary to keep point P rotating around the instantaneous axis of rotation. This is called the Centripetal acceleration.

6. Example: Kinematics of a Human Arm

Question:
Model your arm as two rigid bodies, your forearm of length l and your upper arm of length L, all attached to your body. Suppose your upper arm rotates around a vertical axis with constant angular velocity Ω and your forearm rotates around the axis of your upper arm with constant angular velocity ω. Find the velocity and acceleration of your wrist W with respect to your body.



2.003J Spring 2011: Dynamics and Control I
Massachusetts Institute of Technology
Department of Mechanical Engineering

Derivation of Velocity and Acceleration Equations Using Intermediate Frames
Jason Ku, Feb. 12, 2011

Solution:

First, define our frames of reference. Define reference frame $\hat{\mathrm{B}} = (B, \hat{\mathbf{I}}, \hat{\mathbf{J}}, \hat{\mathbf{K}})$ attached to your body, reference frame $\hat{\mathrm{A}} = (A, \hat{\mathbf{i}}_A, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$ attached to your upper arm, and frame $\hat{\mathrm{F}} = (F, \hat{\mathbf{i}}_F, \hat{\mathbf{j}}_F, \hat{\mathbf{k}}_F)$ attached to your forearm. For convenience, let $\hat{\mathbf{i}}_A$ point in the direction from your shoulder to your elbow and $\hat{\mathbf{i}}_F$ point in the direction from your elbow to your wrist.

Next, we write the given variables with respect to these reference frames. We have:

$$ {}^{\mathrm{A}}\boldsymbol{r}_F = L\,\hat{\mathbf{i}}_A \qquad {}^{\mathrm{F}}\boldsymbol{r}_W = l\,\hat{\mathbf{i}}_F \qquad {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} = \Omega\,\hat{\mathbf{J}} \qquad {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}} = \omega\,\hat{\mathbf{i}}_A $$

Recall the general formula for taking the derivative with respect to frame $\hat{\mathrm{O}}$ of a vector ${}^{\mathrm{A}}\boldsymbol{r}_P$ defined in frame $\hat{\mathrm{A}}$. We will use this relation often in our derivation:

$$ \frac{{}^{\mathrm{O}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_P = {}^{\mathrm{A}}\dot{\boldsymbol{r}}_P + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_P $$

Now, we solve for variable ${}^{\mathrm{B}}\boldsymbol{v}_W$:

$$ \begin{aligned} {}^{\mathrm{B}}\boldsymbol{v}_W = \frac{{}^{\mathrm{B}}d}{dt}\left({}^{\mathrm{B}}\boldsymbol{r}_W\right) = \frac{{}^{\mathrm{B}}d}{dt}\left(\cancelto{0}{{}^{\mathrm{B}}\boldsymbol{r}_A} + {}^{\mathrm{A}}\boldsymbol{r}_F + {}^{\mathrm{F}}\boldsymbol{r}_W\right) &= \frac{{}^{\mathrm{B}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_F + \frac{{}^{\mathrm{B}}d}{dt}{}^{\mathrm{F}}\boldsymbol{r}_W \\ &= (\cancelto{0}{{}^{\mathrm{A}}\dot{\boldsymbol{r}}_F} + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F) + (\cancelto{0}{{}^{\mathrm{F}}\dot{\boldsymbol{r}}_W} + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{F}} \times {}^{\mathrm{F}}\boldsymbol{r}_W) \\ &= \boxed{{}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times {}^{\mathrm{F}}\boldsymbol{r}_W} \end{aligned} $$

Note that we have used the fact that ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{F}} = {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}$. To solve for the acceleration ${}^{\mathrm{B}}\boldsymbol{a}_W$, we simply take another time derivative with respect to the body frame $\hat{\mathrm{B}}$:

$$ \begin{aligned} {}^{\mathrm{B}}\boldsymbol{a}_W &= \frac{{}^{\mathrm{B}}d}{dt}\left({}^{\mathrm{B}}\boldsymbol{v}_W\right) \\ &= \frac{{}^{\mathrm{B}}d}{dt}\left({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times {}^{\mathrm{F}}\boldsymbol{r}_W\right) \\ &= \cancelto{0}{{}^{\mathrm{B}}\dot{\boldsymbol{\omega}}_{\mathrm{A}}} \times {}^{\mathrm{A}}\boldsymbol{r}_F + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times \frac{{}^{\mathrm{B}}d}{dt}{}^{\mathrm{A}}\boldsymbol{r}_F + (\cancelto{0}{{}^{\mathrm{B}}\dot{\boldsymbol{\omega}}_{\mathrm{A}}} + \frac{{}^{\mathrm{B}}d}{dt}{}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times {}^{\mathrm{F}}\boldsymbol{r}_W + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times \frac{{}^{\mathrm{B}}d}{dt}{}^{\mathrm{F}}\boldsymbol{r}_W \\ &= {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times (\cancelto{0}{{}^{\mathrm{A}}\dot{\boldsymbol{r}}_F} + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F) + (\cancelto{0}{{}^{\mathrm{A}}\dot{\boldsymbol{\omega}}_F} + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{\omega}_F) \times {}^{\mathrm{F}}\boldsymbol{r}_W + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times (\cancelto{0}{{}^{\mathrm{F}}\dot{\boldsymbol{r}}_W} + {}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{F}} \times {}^{\mathrm{F}}\boldsymbol{r}_W) \\ &= \boxed{{}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F) + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{\omega}_F) \times {}^{\mathrm{F}}\boldsymbol{r}_W + ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times (({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times {}^{\mathrm{F}}\boldsymbol{r}_W)} \end{aligned} $$

While we have actually done our calculations with respect to two intermediate frames, we can still interpret each term as a specific type of acceleration. The first term ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times ({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{r}_F)$ is the Centripetal acceleration of point A due to the ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}}$ rotation. The second term $({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} \times {}^{\mathrm{A}}\boldsymbol{\omega}_F) \times {}^{\mathrm{F}}\boldsymbol{r}_W$ is the change in the ${}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}$ rotation due to the ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}}$ rotation crossed with vector ${}^{\mathrm{F}}\boldsymbol{r}_W$. This can be interpreted as an effective Eulerian acceleration. The last term $({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times (({}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}} + {}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}) \times {}^{\mathrm{F}}\boldsymbol{r}_W)$ is the Centripetal acceleration of point W due to both the ${}^{\mathrm{B}}\boldsymbol{\omega}_{\mathrm{A}}$ and ${}^{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{F}}$ rotations.

Rigid Body Dynamics

This document contains derivations of Newton's second law and Euler's Equations for rigid bodies derived from Newton's second law for point masses, as well as a definition of the moment of inertia tensor. Students are not responsible for these derivations, but they may be useful in understanding the assumptions and notations used.

Introduction

We accept without proof the linear momentum formulation of Newton's second law for point masses. Since we can directly derive the angular momentum formulation of Newton's second law for point masses directly from the linear formulation, both formulations for point masses yield equivalent sets of equations. A point mass only has three degrees of freedom, so we expect that it should only have three independent equations describing its motion. Rigid bodies on the other hand generally have six degrees of freedom: three in translation and three in rotation. Thus we expect a rigid body to have six independent equations describing its motion. We will see that the linear and angular momentum formulations for rigid bodies derive from two independent assumptions, thus the formulations become independent from each other.

Outer Product, Tensors, and Matrices

For some reference frame $\hat{\text{O}} = (O, \hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$, define vectors $\boldsymbol{a} = a_x\,\hat{\mathbf{i}}_O + a_y\,\hat{\mathbf{j}}_O + a_z\,\hat{\mathbf{k}}_O$ and $\boldsymbol{b} = b_x\,\hat{\mathbf{i}}_O + b_y\,\hat{\mathbf{j}}_O + b_z\,\hat{\mathbf{j}}_O$. The inner product of these vectors is given as the scalar represented by their dot product. Specifically:

$$\boldsymbol{a}\cdot\boldsymbol{b} = (a_x\,\hat{\mathbf{i}}_O + a_y\,\hat{\mathbf{j}}_O + a_z\,\hat{\mathbf{k}}_O)\cdot(b_x\,\hat{\mathbf{i}}_O + b_y\,\hat{\mathbf{j}}_O + b_z\,\hat{\mathbf{j}}_O) = a_xb_x + a_yb_y + a_zb_z$$

In matrix notation, this inner product could be represented as follows:

$$\boldsymbol{a} = {}^{O}\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \quad \boldsymbol{b} = {}^{O}\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} \quad \boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{a}^T\boldsymbol{b} = {}^{O}\begin{bmatrix} a_x & a_y & a_z \end{bmatrix} {}^{O}\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = a_xb_x + a_yb_y + a_zb_z$$

Where $\boldsymbol{a}^T$ represents the transpose of $\boldsymbol{a}$. Note that we use a left superscript O to denote that the elements of this matrix correspond to magnitudes in the $(\hat{\mathbf{i}}_O, \hat{\mathbf{j}}_O, \hat{\mathbf{k}}_O)$ unit vectors. This superscript really refers to frame $\hat{\text{O}}$, not point O, but we write O because it is assumed in this context that we are talking about a frame.

The outer product of two vectors is denoted by the symbol $\otimes$. It is also known as the tensor product or the Kronecker product. It is not the same as the cross product ($\times$). Unlike the inner product, the outer product produces a tensor, not a scalar or vector. The outer product is distributive but not commutative, thus $\boldsymbol{a} \otimes \boldsymbol{b} \neq \boldsymbol{b} \otimes \boldsymbol{a}$. It behaves in the following way:

$$\begin{aligned}
\boldsymbol{a} \otimes \boldsymbol{b} &= (a_x\,\hat{\mathbf{i}}_O + a_y\,\hat{\mathbf{j}}_O + a_z\,\hat{\mathbf{k}}_O) \otimes (b_x\,\hat{\mathbf{i}}_O + b_y\,\hat{\mathbf{j}}_O + b_z\,\hat{\mathbf{j}}_O) \\
&= a_x b_x(\hat{\mathbf{i}}_O \otimes \hat{\mathbf{i}}_O) + a_x b_y(\hat{\mathbf{i}}_O \otimes \hat{\mathbf{j}}_O) + a_x b_z(\hat{\mathbf{i}}_O \otimes \hat{\mathbf{k}}_O) \\
&+ a_y b_x(\hat{\mathbf{j}}_O \otimes \hat{\mathbf{i}}_O) + a_y b_y(\hat{\mathbf{j}}_O \otimes \hat{\mathbf{j}}_O) + a_y b_z(\hat{\mathbf{j}}_O \otimes \hat{\mathbf{k}}_O) \\
&+ a_z b_x(\hat{\mathbf{k}}_O \otimes \hat{\mathbf{i}}_O) + a_z b_y(\hat{\mathbf{k}}_O \otimes \hat{\mathbf{j}}_O) + a_z b_z(\hat{\mathbf{k}}_O \otimes \hat{\mathbf{k}}_O)
\end{aligned}$$

In matrix notation, this outer product would be represented as follows:

$$\boldsymbol{a} \otimes \boldsymbol{b} = \boldsymbol{a}\boldsymbol{b}^T = {}^{O}\!\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} {}^{O}\!\begin{bmatrix} b_x & b_y & b_z \end{bmatrix} = {}^{O}\!\begin{bmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{bmatrix}$$

Outer products multiplied by a third vector can be written as inner products according to the identity $(\boldsymbol{a} \otimes \boldsymbol{b}) \cdot \boldsymbol{c} = (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{a}$. This can be readily verified from the associative property of matrix multiplication:

$$(\boldsymbol{a} \otimes \boldsymbol{b}) \cdot \boldsymbol{c} = (\boldsymbol{a}\boldsymbol{b}^T)\boldsymbol{c} = \boldsymbol{a}(\boldsymbol{b}^T\boldsymbol{c}) = \boldsymbol{a}(\boldsymbol{b} \cdot \boldsymbol{c}) = (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{a} \qquad \checkmark$$

It is important to note that writing vectors as matrices requires that you express it in terms of the unit coordinate vectors of a single frame, and it only makes sense to multiply or add matrices when they are expressed in terms of the same unit vectors. For example, define a frame $\hat{\mathrm{O}} = (O, \hat{\mathbf{I}}, \hat{\mathbf{J}}, \hat{\mathbf{K}})$ and another unit vector $\hat{\mathbf{i}} = \cos\theta\hat{\mathbf{I}} + \sin\theta\hat{\mathbf{J}}$. We could easily define a vector $\boldsymbol{F} = -mg\,\hat{\mathbf{J}} + T\,\hat{\mathbf{i}}$ in terms of any mix of unit coordinates we want. However, to write this vector as a single column vector in matrix notation, we would have to convert to a single basis of unit vectors:

$$\boldsymbol{F} = {}^{O}\!\begin{bmatrix} T\cos\theta \\ T\sin\theta - mg \\ 0 \end{bmatrix} = T\cos\theta\,\hat{\mathbf{I}} + (T\sin\theta - mg)\,\hat{\mathbf{J}}$$

Lastly, the identity matrix $\overline{\mathbb{1}}_3$ is a very special 3x3 matrix. It has the property that for any vector $\boldsymbol{a}$, $\overline{\mathbb{1}}_3\boldsymbol{a} = \boldsymbol{a}$. In addition, its projection onto any set of basis vectors will look the same. If $\overline{\mathbb{1}}_3$ is to be written in terms of the unit vectors of frame $\hat{\mathrm{O}}$, or indeed any frame $\hat{\mathrm{A}}$, it would have the following tensor and matrix representations:

$$\overline{\mathbb{1}}_3 = \hat{\mathbf{I}} \otimes \hat{\mathbf{I}} + \hat{\mathbf{J}} \otimes \hat{\mathbf{J}} + \hat{\mathbf{K}} \otimes \hat{\mathbf{K}} = {}^{O}\!\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \hat{\mathbf{i}}_A \otimes \hat{\mathbf{i}}_A + \hat{\mathbf{j}}_A \otimes \hat{\mathbf{j}}_A + \hat{\mathbf{k}}_A \otimes \hat{\mathbf{k}}_A = {}^{A}\!\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Linear Momentum for Rigid Bodies: Extension of Newton's Second Law

Define a fixed inertial reference frame $\hat{\text{O}} = (O, \hat{\mathbf{I}}, \hat{\mathbf{J}}, \hat{\mathbf{K}})$. Let us also define a discrete set of N point masses as a rigid body M with $M = \sum_{i=1}^{N} m_i$. First, let us derive an expression relating the linear momentum of the rigid body to the forces acting on it. For each point mass in the body, the linear formulation of Newton's second law applies:

$$\sum \boldsymbol{F}_i = \frac{{}^{\text{O}}d}{dt}\left({}^{\text{O}}\boldsymbol{p}_i\right) \qquad \text{for } i = 1 \to N$$

Where $\sum \boldsymbol{F}_i$ is the sum of all the forces acting on point mass m_i. Let us sum together all of these equations for all N point masses:

$$\sum_{i=1}^{N} \sum \boldsymbol{F}_i = \sum_{i=1}^{N} \frac{{}^{\text{O}}d}{dt}\left({}^{\text{O}}\boldsymbol{p}_i\right) \tag{1}$$

Let us first examine the left side of equation (1). Every force acting on each point mass must either be an internal force from another point mass in the extended body, or an external force. Define $\boldsymbol{F}_{ij}$ to be the internal force that point mass m_j exerts on point mass m_i, and $\sum \boldsymbol{F}_M$ to be the sum of external forces on the collection of particles M. Then:

$$\sum_{i=1}^{N} \sum \boldsymbol{F}_i = \underbrace{\left(\sum \boldsymbol{F}_M\right)}_{external} + \underbrace{\left(\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{F}_{ij}\right)}_{internal}$$

Here we must use an independent assertion to proceed. We will exploit Newton's third law such that the reaction forces between interacting point masses in the body must be equal and opposite, or $\boldsymbol{F}_{ij} \equiv -\boldsymbol{F}_{ji}$. Thus we have that:

$$\underbrace{\left(\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{F}_{ij}\right)}_{internal} = 0$$

Let us assume the rigid body's mass is constant in time. Because the derivative operator is distributive, the right side of equation (1) becomes:

$$\sum_{i=1}^{N} \frac{{}^{\text{O}}d}{dt}\left({}^{\text{O}}\boldsymbol{p}_i\right) = \frac{{}^{\text{O}}d}{dt}\left(\sum_{i=1}^{N} m_i {}^{\text{O}}\boldsymbol{v}_i\right) = M\frac{{}^{\text{O}}d}{dt}\left[\frac{{}^{\text{O}}d}{dt}\left(\frac{1}{M}\sum_{i=1}^{N} m_i {}^{\text{O}}\boldsymbol{r}_i\right)\right] = M {}^{\text{O}}\boldsymbol{a}_C$$

Here we have used the definition of the center of mass C of a discrete system or, in the limiting case, a continuum with a density as a function of position $\rho({}^{\text{O}}\boldsymbol{r}_V)$ over volume V:

$$ {}^{\mathrm{O}}\boldsymbol{r}_C = \frac{1}{M}\sum_{i=1}^{N} m_i {}^{\mathrm{O}}\boldsymbol{r}_i \qquad \text{or} \qquad {}^{\mathrm{O}}\boldsymbol{r}_C = \frac{1}{M}\iiint_V \rho({}^{\mathrm{O}}\boldsymbol{r}_V){}^{\mathrm{O}}\boldsymbol{r}_V\, dV \quad \text{with} \quad M = \iiint_V \rho({}^{\mathrm{O}}\boldsymbol{r}_V)\, dV $$

Plugging back into equation (1) yields the linear momentum formulation of Newton's second law for a rigid body relating the sum of the external forces acting on the body to the acceleration of the body's center of mass.

$$ \boxed{\sum \boldsymbol{F}_M = M {}^{\mathrm{O}}\boldsymbol{a}_C} $$

Angular Momentum for Rigid Bodies: Euler's Equations

Now let us derive an expression relating the angular momentum of the rigid body to the torques acting on it. Define a reference frame $\hat{\mathrm{B}} = (B, \hat{\mathbf{i}}_B, \hat{\mathbf{j}}_B, \hat{\mathbf{k}}_B)$ in which the rigid body is stationary, i.e. a frame moving with the body. For each point mass in the body, the angular momentum formulation of Newton's second law applies:

$$ \sum \boldsymbol{\tau}_i^B = \frac{{}^{\mathrm{O}}d}{dt}\left({}^{O}\boldsymbol{h}_i^B\right) + {}^{\mathrm{O}}\boldsymbol{v}_B \times {}^{\mathrm{O}}\boldsymbol{p}_i \qquad\qquad \text{for } i = 1 \to N $$

Where $\sum \boldsymbol{\tau}_i^B$ is the sum of the torques about point B acting on point mass m_i.

Let us sum these equations over all N point masses:

$$ \sum_{i=1}^{N}\left({}^{\mathrm{B}}\boldsymbol{r}_i \times \sum \boldsymbol{F}_i\right) = \sum_{i=1}^{N}\frac{{}^{\mathrm{O}}d}{dt}\left({}^{O}\boldsymbol{h}_i^B\right) + \sum_{i=1}^{N}{}^{\mathrm{O}}\boldsymbol{v}_B \times {}^{\mathrm{O}}\boldsymbol{p}_i \tag{2} $$

Let us first examine the left side of equation (2). Again, every force acting on each point mass will either be an internal force from another point mass in the extended body or an external force acting on the body. Define $\sum \boldsymbol{\tau}_M^B$ to be the sum of external torques on the collection of particles M about point B. Then:

$$ \sum_{i=1}^{N}\left({}^{\mathrm{B}}\boldsymbol{r}_i \times \sum \boldsymbol{F}_i\right) = \underbrace{\left(\sum \boldsymbol{\tau}_M^B\right)}_{external} + \underbrace{\left[\sum_{i=1}^{N}\left({}^{\mathrm{B}}\boldsymbol{r}_i \times \sum_{j=1}^{N}\boldsymbol{F}_{m_{ij}}\right)\right]}_{internal} $$

Here again we must use an independent assertion to proceed. We could just define a rigid body such that the sum of the torques on the body due to internal forces sum to zero. Alternatively, we can model the point masses in the rigid body such that they cannot move relative to each other, thus the sum of the internal forces can do no work on the rigid body. Define W_{ij} to be the work that point mass m_j exerts on point mass m_i. Then:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij} = 0 = \sum_{i=1}^{N}\sum_{j=1}^{N}\int \boldsymbol{F}_{ij}\cdot {}^{\mathrm{O}}\boldsymbol{d}\boldsymbol{r}_i = \int \sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{F}_{ij}\cdot {}^{\mathrm{O}}\boldsymbol{v}_i\,dt$$

This must hold true for any time interval, thus, for any frame $\hat{\mathrm{B}}$ in which every point mass m_i is stationary:

$$0 = \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{F}_{ij}\cdot {}^{\mathrm{O}}\boldsymbol{v}_i = \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{F}_{ij}\cdot \frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{O}}\boldsymbol{r}_B + {}^{\mathrm{B}}\boldsymbol{r}_i) = \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{F}_{ij}\cdot({}^{\mathrm{O}}\boldsymbol{v}_B + \cancelto{0}{{}^{\mathrm{B}}\boldsymbol{v}_i} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\times{}^{\mathrm{B}}\boldsymbol{r}_i)$$

$$= {}^{\mathrm{O}}\boldsymbol{v}_B\cdot\cancelto{0}{\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{F}_{ij}} + \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{F}_{ij}\cdot({}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\times{}^{\mathrm{B}}\boldsymbol{r}_i) = {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\cdot\sum_{i=1}^{N}\left({}^{\mathrm{B}}\boldsymbol{r}_i\times\sum_{j=1}^{N}\boldsymbol{F}_{ij}\right)$$

Since this equation must hold true for any angular velocity ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}$ of the rigid body:

$$\underbrace{\left[\sum_{i=1}^{N}\left({}^{\mathrm{B}}\boldsymbol{r}_i\times\sum_{j=1}^{N}\boldsymbol{F}_{ij}\right)\right]}_{internal} = 0$$

Let us assume the rigid body's mass is constant in time. Because the derivative operator is distributive, the second term on the right side of equation (2) becomes:

$${}^{\mathrm{O}}\boldsymbol{v}_B\times\sum_{i=1}^{N}{}^{\mathrm{O}}\boldsymbol{p}_i = {}^{\mathrm{O}}\boldsymbol{v}_B\times\frac{{}^{\mathrm{O}}d}{dt}\sum_{i=1}^{N}m_i\,{}^{\mathrm{O}}\boldsymbol{r}_i = M\,{}^{\mathrm{O}}\boldsymbol{v}_B\times{}^{\mathrm{O}}\boldsymbol{v}_C$$

For the first term on the right side of equation (2), let us analyze the angular momentum ${}^{O}\boldsymbol{h}_i^B$. Recall the vector identity $\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}$. Applying this identity:

$$\sum_{i=1}^{N}{}^{O}\boldsymbol{h}_i^B = \sum_{i=1}^{N}m_i\,{}^{\mathrm{B}}\boldsymbol{r}_i\times{}^{\mathrm{O}}\boldsymbol{v}_i = \sum_{i=1}^{N}m_i\,{}^{\mathrm{B}}\boldsymbol{r}_i\times({}^{\mathrm{O}}\boldsymbol{v}_B + \cancelto{0}{{}^{\mathrm{B}}\boldsymbol{v}_i} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\times{}^{\mathrm{B}}\boldsymbol{r}_i)$$

$$= M\,{}^{\mathrm{B}}\boldsymbol{r}_C\times{}^{\mathrm{O}}\boldsymbol{v}_B + \sum_{i=1}^{N}m_i\left[({}^{\mathrm{B}}\boldsymbol{r}_i\cdot{}^{\mathrm{B}}\boldsymbol{r}_i){}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}} - ({}^{\mathrm{B}}\boldsymbol{r}_i\cdot{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}){}^{\mathrm{B}}\boldsymbol{r}_i\right]$$

We will utilize two more vector/matrix identities: $\boldsymbol{c} = \mathbb{1}_3\boldsymbol{c}$, where $\mathbb{1}_3$ is the 3x3 identity matrix, and $(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c} = (\boldsymbol{c}\otimes\boldsymbol{a})\boldsymbol{b}$, where $\otimes$ is the outer product. Applying these rules, we get:

$$\sum_{i=1}^{N} {}^{O}\boldsymbol{h}_i^B = M\,{}^{\mathrm{B}}\boldsymbol{r}_C \times {}^{\mathrm{O}}\boldsymbol{v}_B + \sum_{i=1}^{N} m_i \left[({}^{\mathrm{B}}\boldsymbol{r}_i \cdot {}^{\mathrm{B}}\boldsymbol{r}_i)\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_i \otimes {}^{\mathrm{B}}\boldsymbol{r}_i)\right] {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}} = M\,{}^{\mathrm{B}}\boldsymbol{r}_C \times {}^{\mathrm{O}}\boldsymbol{v}_B + \bar{\mathbf{I}}_M^B\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}$$

This is how we define the moment of inertia $\bar{\mathbf{I}}_M^B$ of a rigid body mass M about pivot point B. This is a 3x3 tensor. To get a better sense of this tensor, let us express it in terms of its components in the $\hat{\mathrm{B}}$ frame, with ${}^{\mathrm{B}}\boldsymbol{r}_i = x_i\,\hat{\mathbf{i}}_B + y_i\,\hat{\mathbf{j}}_B + z_i\,\hat{\mathbf{k}}_B$. Thus:

$$\begin{aligned}
\bar{\mathbf{I}}_M^B &= \sum_{i=1}^{N} m_i \left[({}^{\mathrm{B}}\boldsymbol{r}_i \cdot {}^{\mathrm{B}}\boldsymbol{r}_i)\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_i \otimes {}^{\mathrm{B}}\boldsymbol{r}_i)\right] \\
&= \sum_{i=1}^{N} m_i \left((x_i^2 + y_i^2 + z_i^2)\ {}^{B}\!\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - {}^{B}\!\begin{bmatrix} x_i^2 & x_i y_i & x_i z_i \\ y_i x_i & y_i^2 & y_i z_i \\ z_i x_i & z_i y_i & z_i^2 \end{bmatrix} \right) \\
&= \sum_{i=1}^{N} m_i\ {}^{B}\!\begin{bmatrix} y_i^2 + z_i^2 & -x_i y_i & -x_i z_i \\ -y_i x_i & z_i^2 + x_i^2 & -y_i z_i \\ -z_i x_i & -z_i y_i & x_i^2 + y_i^2 \end{bmatrix}
\end{aligned}$$

In the continuous case:

$$\begin{aligned}
\bar{\mathbf{I}}_M^B &= \iiint_V \rho({}^{\mathrm{B}}\boldsymbol{r}_{dV}) \left[({}^{\mathrm{B}}\boldsymbol{r}_{dV} \cdot {}^{\mathrm{B}}\boldsymbol{r}_{dV})\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_{dV} \otimes {}^{\mathrm{B}}\boldsymbol{r}_{dV})\right] dV \\
&= \iiint_V \rho(x,y,z)\ {}^{B}\!\begin{bmatrix} y^2 + z^2 & -xy & -xz \\ -yx & z^2 + x^2 & -yz \\ -zx & -zy & x^2 + y^2 \end{bmatrix} dx\,dy\,dz
\end{aligned}$$

The diagonal terms of this matrix are called the moments of inertia while the off-diagonal terms are called the products of inertia. Here, these matrices have been expressed in terms of the unit coordinates of the $\hat{\mathrm{B}}$ frame, but we could have expressed in terms of the unit coordinates of <u>any</u> frame. Plugging back into equation (2) yields Euler's Equations for a rigid body. For a rigid body we assume that the moment of inertia tensor is constant in time with respect to frame $\hat{\mathrm{B}}$ in which the body is stationary.

$$\begin{aligned}
\sum \boldsymbol{\tau}_M^B &= \frac{{}^{\mathrm{O}}d}{dt}\left(M\,{}^{\mathrm{B}}\boldsymbol{r}_C \times {}^{\mathrm{O}}\boldsymbol{v}_B + \bar{\mathbf{I}}_M^B\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\right) + M\,{}^{\mathrm{O}}\boldsymbol{v}_B \times {}^{\mathrm{O}}\boldsymbol{v}_C \\
&= M\left[\frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{B}}\boldsymbol{r}_C \times {}^{\mathrm{O}}\boldsymbol{v}_B) + {}^{\mathrm{O}}\boldsymbol{v}_B \times {}^{\mathrm{O}}\boldsymbol{v}_C\right] + \bar{\mathbf{I}}_M^B\,{}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{B}} + \left(\cancelto{0}{\frac{{}^{\mathrm{B}}d}{dt}\bar{\mathbf{I}}_M^B} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}} \times \bar{\mathbf{I}}_M^B\right){}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}
\end{aligned}$$

$$\boxed{\sum \boldsymbol{\tau}_M^B = \bar{\mathbf{I}}_M^B\,{}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{B}} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}} \times \left(\bar{\mathbf{I}}_M^B\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\right) + M\left[\frac{{}^{\mathrm{O}}d}{dt}({}^{\mathrm{B}}\boldsymbol{r}_C \times {}^{\mathrm{O}}\boldsymbol{v}_B) + {}^{\mathrm{O}}\boldsymbol{v}_B \times {}^{\mathrm{O}}\boldsymbol{v}_C\right]}$$

Note that if we choose $B \equiv C$ <u>or</u> B fixed in frame Ô, then the last term goes to zero, and we get a simplified version of Euler's Equations:

$$\boxed{\sum \boldsymbol{\tau}_M^B = \bar{\mathbf{I}}_M^B \, {}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{B}} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}} \times \left(\bar{\mathbf{I}}_M^B \, {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{B}}\right)}$$

Parallel Axis Theorem

From the definition of the general moment of inertia tensor $\bar{\mathbf{I}}_M^B$, we can readily break up the vector ${}^{\mathrm{B}}\boldsymbol{r}_i$ up into two components such that ${}^{\mathrm{B}}\boldsymbol{r}_i = {}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i$ where C is the center of mass of the rigid body.

$$\bar{\mathbf{I}}_M^B = \sum_{i=1}^N m_i \left[(({}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i) \cdot ({}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i))\bar{\mathbb{1}}_3 - (({}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i) \otimes ({}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i)) \right]$$

Exploiting the distributive property:

$$\begin{aligned}
\bar{\mathbf{I}}_M^B &= \sum_{i=1}^N m_i \left[({}^{\mathrm{B}}\boldsymbol{r}_C \cdot {}^{\mathrm{B}}\boldsymbol{r}_C + 2{}^{\mathrm{B}}\boldsymbol{r}_C \cdot {}^{\mathrm{C}}\boldsymbol{r}_i + {}^{\mathrm{C}}\boldsymbol{r}_i \cdot {}^{\mathrm{C}}\boldsymbol{r}_i))\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_C \otimes {}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{B}}\boldsymbol{r}_C \otimes {}^{\mathrm{C}}\boldsymbol{r}_i + {}^{\mathrm{C}}\boldsymbol{r}_i \otimes {}^{\mathrm{B}}\boldsymbol{r}_C + {}^{\mathrm{C}}\boldsymbol{r}_i \otimes {}^{\mathrm{C}}\boldsymbol{r}_i)) \right] \\
&= \sum_{i=1}^N m_i \left[({}^{\mathrm{B}}\boldsymbol{r}_C \cdot {}^{\mathrm{B}}\boldsymbol{r}_C)\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_C \otimes {}^{\mathrm{B}}\boldsymbol{r}_C) \right] + \sum_{i=1}^N m_i \left[({}^{\mathrm{C}}\boldsymbol{r}_i \cdot {}^{\mathrm{C}}\boldsymbol{r}_i)\bar{\mathbb{1}}_3 - ({}^{\mathrm{C}}\boldsymbol{r}_i \otimes {}^{\mathrm{C}}\boldsymbol{r}_i) \right] \\
&\qquad + 2{}^{\mathrm{B}}\boldsymbol{r}_C \cdot \sum_{i=1}^N m_i {}^{\mathrm{C}}\boldsymbol{r}_i + {}^{\mathrm{B}}\boldsymbol{r}_C \otimes \sum_{i=1}^N m_i {}^{\mathrm{C}}\boldsymbol{r}_i + \sum_{i=1}^N m_i {}^{\mathrm{C}}\boldsymbol{r}_i \otimes {}^{\mathrm{B}}\boldsymbol{r}_C
\end{aligned}$$

But $\sum_{i=1}^N m_i {}^{\mathrm{C}}\boldsymbol{r}_i = M {}^{\mathrm{C}}\boldsymbol{r}_C = \mathbf{0}$, thus, if ${}^{\mathrm{B}}\boldsymbol{r}_C = a\,\hat{\mathbf{i}}_B + b\,\hat{\mathbf{j}}_B + c\,\hat{\mathbf{k}}_B$, we have:

$$\boxed{\bar{\mathbf{I}}_M^B = \bar{\mathbf{I}}_M^C + M \left[({}^{\mathrm{B}}\boldsymbol{r}_C \cdot {}^{\mathrm{B}}\boldsymbol{r}_C)\bar{\mathbb{1}}_3 - ({}^{\mathrm{B}}\boldsymbol{r}_C \otimes {}^{\mathrm{B}}\boldsymbol{r}_C) \right]}$$

$$\bar{\mathbf{I}}_M^B = \bar{\mathbf{I}}_M^C + M \;{}^{B}\!\begin{bmatrix} b^2 + c^2 & -ab & -ac \\ -ba & c^2 + a^2 & -bc \\ -ca & -cb & a^2 + b^2 \end{bmatrix}$$

Symmetry and Products of Inertia

If a rigid body has a plane of symmetry about the $C\,\hat{\mathbf{j}}_B\,\hat{\mathbf{k}}_B$ plane, then $\rho(x, y, z) = \rho(-x, y, z)$. This tells us something about the products of inertia in the inertia tensor $\bar{\mathbf{I}}_M^C$. For example, let us examine the $\hat{\mathbf{i}}_B \otimes \hat{\mathbf{j}}_B$

term in this tensor:

$$\begin{aligned}
\bar{\mathbf{I}}_M^C \cdot (\hat{\mathbf{i}}_B \otimes \hat{\mathbf{j}}_B) = \mathrm{I}_{xy}^C &= \iiint_{-\infty}^{\infty} \rho(x,y,z)(-xy)dx\,dy\,dz \\
&= \iint_{-\infty}^{\infty} \left[\int_0^{\infty} \rho(x,y,z)(-xy)dx + \int_{-\infty}^{0} \rho(x,y,z)(-xy)dx \right] dy\,dz \\
&= \iint_{-\infty}^{\infty} \left[\int_0^{\infty} \rho(x,y,z)(-xy)dx + \int_{\infty}^{0} \rho(-x,y,z)(xy)(-dx) \right] dy\,dz \\
&= \iint_{-\infty}^{\infty} \left[\int_0^{\infty} \rho(x,y,z)(-xy)dx - \int_{0}^{\infty} \rho(-x,y,z)(-xy)dx \right] dy\,dz \\
&= 0
\end{aligned}$$

Conceptually, every point on the positive $\hat{\mathbf{i}}_B$ side of the rigid body has an equivalent density on the negative $\hat{\mathbf{i}}_B$ side, so the integral with respect to the x variable must sum to zero. Thus all the products of inertia containing an x variable, i.e. I_{xy}^C, I_{yx}^C, I_{xz}^C, and I_{zx}^C must all be zero. Also notice that if the any two of the $C\,\hat{\mathbf{i}}_B\,\hat{\mathbf{j}}_B$, $C\,\hat{\mathbf{j}}_B\,\hat{\mathbf{k}}_B$, and $C\,\hat{\mathbf{k}}_B\,\hat{\mathbf{i}}_B$ planes are planes of symmetry, all products of inertia will be zero.

Principal Axes

We have seen that the inertia tensor $\bar{\mathbf{I}}_M^C$ about a rigid body's center of mass C can be described as a matrix in terms of the unit basis coordinates of some reference frame. Linear algebra tells us that because the inertia tensor is both real and symmetric, there must exist some choice of basis coordinates such that $\bar{\mathbf{I}}_M^C$ is diagonal. We call this set of basis coordinates the body's principal axes. The key property of principal axes is that multiplying $\bar{\mathbf{I}}_M^C$ by a principal direction results in a vector pointing in the same direction:

$$\bar{\mathbf{I}}_M^C \cdot \hat{\mathbf{i}}_C = I_{xx}\,\hat{\mathbf{i}}_C$$

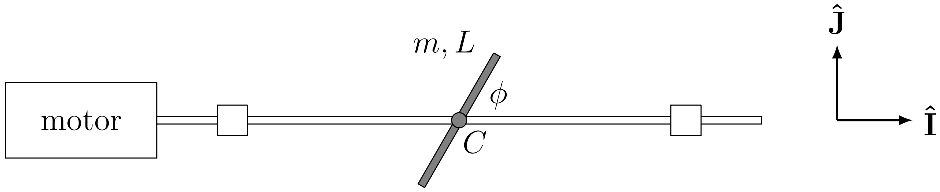
Here, we call the principal direction $\hat{\mathbf{i}}_C$ an eigenvector of the matrix $\bar{\mathbf{I}}_M^C$, and call the principal moment of inertia I_{xx} the corresponding eigenvalue. We can exploit this property to find the principal directions of a body. Suppose we calculate the tensor components of $\bar{\mathbf{I}}_M^C$ in terms of a frame whose unit vectors are not the principal directions. If we can find a vector $\hat{\mathbf{i}}_C$ such that $\bar{\mathbf{I}}_M^C \cdot \hat{\mathbf{i}}_C = I\,\hat{\mathbf{i}}_C$ for some constant I, we will have found a principal direction. Typically we would parameterize $\hat{\mathbf{i}}_C$ in terms of angles of rotation, perform the $\bar{\mathbf{I}}_M^C \cdot \hat{\mathbf{i}}$ multiplication, set the $\hat{\mathbf{j}}_C$ and $\hat{\mathbf{k}}_C$ terms to zero, and solve for the angles of rotation in order to find the principal directions.

Note that if a rigid body has certain symmetries, finding the principal axes may be quite easy to find, but may not be unique. For example, if a rigid body is rotationally symmetric about a single axis, any basis set of unit vectors including the axial direction will be a set of principal axes. This fact can be proved from the symmetry arguments above, noting that all products of inertia in terms of such a basis must cancel to zero. Also, for any rigid body with a plane of symmetry, the normal direction to the plane will be a principal direction. Typically, we will not need to solve for the principal directions as we can just look up the moment of inertia for common objects, which will be given in terms of their principal directions.

If we write Euler's Equations in terms of a frame $\hat{\mathrm{C}} = (C, \hat{\mathbf{i}}_C, \hat{\mathbf{j}}_C, \hat{\mathbf{k}}_C)$ associated with the principal axes of the body such that $\bar{\mathbf{I}}_M^C = I_{xx}\,\hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C + I_{yy}\,\hat{\mathbf{j}}_C \otimes \hat{\mathbf{j}}_C + I_{zz}\,\hat{\mathbf{k}}_C \otimes \hat{\mathbf{k}}_C$, ${}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{C}} = \omega_x\,\hat{\mathbf{i}}_C + \omega_y\,\hat{\mathbf{j}}_C + \omega_z\,\hat{\mathbf{k}}_C$, and $\boldsymbol{\tau}_M^C = \tau_x\,\hat{\mathbf{i}}_C + \tau_y\,\hat{\mathbf{j}}_C + \tau_z\,\hat{\mathbf{k}}_C$, Euler's Equations can be written as the following three scalar equations:

$$\tau_x = I_{xx}\omega_x + \omega_y\omega_z(I_{yy} - I_{zz})$$
$$\tau_y = I_{yy}\omega_y + \omega_z\omega_x(I_{zz} - I_{xx})$$
$$\tau_y = I_{zz}\omega_z + \omega_x\omega_y(I_{xx} - I_{yy})$$

Example



A motor rotates a shaft of negligible mass at a constant angular velocity $\dot{\theta}\,\hat{\mathbf{I}}$. The shaft is held on two bearings. Halfway between these bearings, a bar of uniform density with mass m and length L is rigidly fixed to the shaft at an angle ϕ from the axis of the shaft. Determine the torque the bearings must exert on the shaft-bar system to keep the system rotating at constant velocity.

Solution

Define ground reference frame $\hat{\mathrm{O}} = (O \equiv C, \hat{\mathbf{I}}, \hat{\mathbf{J}}, \hat{\mathbf{K}})$, intermediate frame $\hat{\mathrm{A}} = (A \equiv C, \hat{\mathbf{i}}_A \equiv \hat{\mathbf{I}}, \hat{\mathbf{j}}_A, \hat{\mathbf{k}}_A)$ rotating and aligned with the massless shaft such that ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{A}} = \dot{\theta}\,\hat{\mathbf{I}}$ and $\hat{\mathbf{J}} \times \hat{\mathbf{j}}_A = \sin\theta\,\hat{\mathbf{I}}$, and intermediate frame $\hat{\mathrm{C}} = (C, \hat{\mathbf{i}}_C, \hat{\mathbf{j}}_C, \hat{\mathbf{k}}_C \equiv \hat{\mathbf{k}}_A)$ rotating and aligned with the bar such that ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} = \dot{\theta}\,\hat{\mathbf{I}}$ and $\hat{\mathbf{i}}_A \times \hat{\mathbf{i}}_C = \sin\phi\,\hat{\mathbf{k}}_A$. Let us write Euler's Equations for the sum of the torques about the center of mass, point C:

$$\sum \boldsymbol{\tau}_m^C = \bar{\mathbf{I}}_m^C \cancelto{0}{{}^{\mathrm{O}}\dot{\boldsymbol{\omega}}_{\mathrm{C}}} + {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} \times \left(\bar{\mathbf{I}}_m^C\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}\right) = {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} \times \left(\bar{\mathbf{I}}_m^C\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}\right)$$

Note that we can use the simplified Euler Equations because we are summing torques about the center of mass C, which also happens to be fixed in ground frame $\hat{\mathrm{O}}$. The left side of the equation is what we are looking for, while the only thing on the right side we don't know is the moment of inertia tensor $\bar{\mathbf{I}}_m^C$.

$$\bar{\mathbf{I}}_m^C = \iiint_V \rho({}^{\mathrm{C}}\boldsymbol{r}_{dV}) \left[({}^{\mathrm{C}}\boldsymbol{r}_{dV} \cdot {}^{\mathrm{C}}\boldsymbol{r}_{dV})\bar{\mathbb{1}}_3 - ({}^{\mathrm{C}}\boldsymbol{r}_{dV} \otimes {}^{\mathrm{C}}\boldsymbol{r}_{dV})\right] dV$$

Let us express this tensor first in terms of frame $\hat{\mathrm{C}}$. The mass per unit length of the rod is a constant $\frac{m}{L}$. We must integrate along the length of the rod, which we can parameterize as ${}^{\mathrm{C}}\boldsymbol{r}_{dV} = x\,\hat{\mathbf{i}}_C$:

$$\bar{\mathbf{I}}_m^C = \int_{-L/2}^{L/2} \frac{m}{L}\left[(x^2\,\hat{\mathbf{i}}_C \cdot \hat{\mathbf{i}}_C)\bar{\mathbb{1}}_3 - (x^2\,\hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C)\right] dx = \frac{m}{L}\int_{-L/2}^{L/2} x^2 dx \left(\bar{\mathbb{1}}_3 - \hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C\right) = \frac{mL^2}{12}\left(\bar{\mathbb{1}}_3 - \hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C\right)$$

We can write this in terms of both tensor and matrix notation:

$$\begin{aligned}\bar{\mathbf{I}}_m^C &= \frac{mL^2}{12}\left[(\hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C + \hat{\mathbf{j}}_C \otimes \hat{\mathbf{j}}_C + \hat{\mathbf{k}}_C \otimes \hat{\mathbf{k}}_C) - \hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C\right] = \frac{mL^2}{12}\left(\hat{\mathbf{j}}_C \otimes \hat{\mathbf{j}}_C + \hat{\mathbf{k}}_C \otimes \hat{\mathbf{k}}_C\right) \\ &= \frac{mL^2}{12}\left({}^{C}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - {}^{C}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \frac{mL^2}{12}\,{}^{C}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\end{aligned}$$

Note that this tensor is diagonal, which means we happened to project the moment of inertia tensor onto a set of principal axes. Note that one rarely needs to calculate moments of inertia explicitly as the inertia tensor along the principal axes for most common shapes can be readily looked up in tables. Plugging into Euler's Equations, we find:

$$\begin{aligned}\sum \boldsymbol{\tau}_m^C &= \dot{\theta}\,\hat{\mathbf{I}} \times \left[\frac{mL^2}{12}\left(\hat{\mathbf{j}}_C \otimes \hat{\mathbf{j}}_C + \hat{\mathbf{k}}_C \otimes \hat{\mathbf{k}}_C\right) \cdot \dot{\theta}\,\hat{\mathbf{I}}\right] = \frac{mL^2}{12}\dot{\theta}^2\,\hat{\mathbf{I}} \times [(\hat{\mathbf{j}}_C \cdot \hat{\mathbf{I}})\hat{\mathbf{j}}_C + \cancelto{0}{(\hat{\mathbf{k}}_C \cdot \hat{\mathbf{I}})}\hat{\mathbf{k}}_C] \\ &= \frac{mL^2}{12}\dot{\theta}^2(-\sin\phi\,\hat{\mathbf{I}} \times \hat{\mathbf{j}}_C) = \boxed{-\frac{mL^2}{12}\dot{\theta}^2 \sin\phi\cos\phi\,\hat{\mathbf{k}}_A}\end{aligned}$$

If we wanted to use matrix notation, we would need to either convert $\bar{\mathbf{I}}_m^C$ to the ground reference frame coordinates or convert ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}$ to the unit coordinates of frame $\hat{\mathrm{C}}$. Both methods yield the same answer as the tensor representation, though writing the inertia tensor in the ground coordinates is quite tedious.

Converting $\bar{\mathbf{I}}_m^C$ to the ground frame coordinates:

$$\bar{\mathbf{I}}_m^C = \frac{mL^2}{12}\left(\bar{\mathbb{1}}_3 - \hat{\mathbf{i}}_C \otimes \hat{\mathbf{i}}_C\right) = \frac{mL^2}{12}\left(\bar{\mathbb{1}}_3 - (\cos\phi\,\hat{\mathbf{i}}_A + \sin\phi\,\hat{\mathbf{j}}_A) \otimes (\cos\phi\,\hat{\mathbf{i}}_A + \sin\phi\,\hat{\mathbf{j}}_A)\right)$$

$$= \frac{mL^2}{12}\left(\bar{\mathbb{1}}_3 - [\cos\phi\,\hat{\mathbf{I}} + \sin\phi(\cos\theta\hat{\mathbf{J}} + \sin\theta\hat{\mathbf{K}})] \otimes [\cos\phi\,\hat{\mathbf{I}} + \sin\phi(\cos\theta\hat{\mathbf{J}} + \sin\theta\hat{\mathbf{K}})]\right)$$

$$= \frac{mL^2}{12}\left({}^{O}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - {}^{O}\begin{bmatrix} \cos^2\phi & \sin\phi\cos\phi\cos\theta & \sin\phi\cos\phi\sin\theta \\ \sin\phi\cos\phi\cos\theta & \sin^2\phi\cos^2\theta & \sin^2\phi\sin\theta\cos\theta \\ \sin\phi\cos\phi\sin\theta & \sin^2\phi\sin\theta\cos\theta & \sin^2\phi\sin^2\theta \end{bmatrix}\right)$$

$$= \frac{mL^2}{12}\left({}^{O}\begin{bmatrix} 1-\cos^2\phi & -\sin\phi\cos\phi\cos\theta & -\sin\phi\cos\phi\sin\theta \\ -\sin\phi\cos\phi\cos\theta & 1-\sin^2\phi\cos^2\theta & -\sin^2\phi\sin\theta\cos\theta \\ -\sin\phi\cos\phi\sin\theta & -\sin^2\phi\sin\theta\cos\theta & 1-\sin^2\phi\sin^2\theta \end{bmatrix}\right)$$

$$\sum \boldsymbol{\tau}_m^C = {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} \times \left(\bar{\mathbf{I}}_m^C\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}\right) = \dot{\theta}\;{}^{O}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \frac{mL^2}{12}\dot{\theta}\;{}^{O}\begin{bmatrix} 1-\cos^2\phi \\ -\sin\phi\cos\phi\cos\theta \\ -\sin\phi\cos\phi\sin\theta \end{bmatrix}$$

$$= \frac{mL^2}{12}\dot{\theta}^2\sin\phi\cos\phi(\sin\theta\hat{\mathbf{J}} - \cos\theta\hat{\mathbf{K}}) = \boxed{-\frac{mL^2}{12}\dot{\theta}^2\sin\phi\cos\phi\,\hat{\mathbf{k}}_A}$$

In general, we will almost always prefer to keep the inertia tensor in terms of the principal directions, and just convert the angular velocity to match. Converting ${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}$ to the unit coordinates of frame $\hat{\mathrm{C}}$:

$${}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} = \dot{\theta}\,\hat{\mathbf{I}} = \dot{\theta}\,\hat{\mathbf{i}}_A = \dot{\theta}(\cos\phi\,\hat{\mathbf{i}}_C - \sin\phi\,\hat{\mathbf{j}}_C) = \dot{\theta}\;{}^{C}\begin{bmatrix} \cos\phi \\ -\sin\phi \\ 0 \end{bmatrix}$$

$$\sum \boldsymbol{\tau}_m^C = {}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}} \times \left(\bar{\mathbf{I}}_m^C\,{}^{\mathrm{O}}\boldsymbol{\omega}_{\mathrm{C}}\right) = \dot{\theta}\;{}^{C}\begin{bmatrix} \cos\phi \\ -\sin\phi \\ 0 \end{bmatrix} \times \left(\frac{mL^2}{12}\;{}^{O}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\dot{\theta}\;{}^{C}\begin{bmatrix} \cos\phi \\ -\sin\phi \\ 0 \end{bmatrix}\right)$$

$$= \frac{mL^2}{12}\dot{\theta}^2\;{}^{C}\begin{bmatrix} \cos\phi \\ -\sin\phi \\ 0 \end{bmatrix} \times {}^{C}\begin{bmatrix} 0 \\ -\sin\phi \\ 0 \end{bmatrix} = \boxed{-\frac{mL^2}{12}\dot{\theta}^2\sin\phi\cos\phi\,\hat{\mathbf{k}}_A}$$

Note that when $\phi = 0$ or $\phi = \frac{\pi}{2}$, no torque is required. However, for all other values of ϕ, the torque on the bearings will be non-zero, constantly and periodically changing direction with time which can cause significant rattling, fatigue, and damage.