A Complete Derivation of the Newton-Euler Equations

A Document

Jake Panikulam
Main Street Autonomy

July 28, 2023

1 Introduction

For point particles, we know $\mathbf{F} = m\mathbf{a}$, Newton's second law. Point particles can be fully described by their position, their velocity and their mass. Rigid bodies can be thought of as an arrangement of infinitely many point particles. It is hard to work with infinite collections of particles when we are computing motions. Instead, we prefer to summarize describe the collection of particles as having position and **rotation**. Rigid bodies can then be fully described by their position, velocity, mass and newcomers rotation, angular velocity and rotational inertia.

The "Newton-Euler" equations are the equivalent of $\mathbf{F} = m\mathbf{a}$ for rigid bodies. $\mathbf{F} = m\mathbf{a}$ applies to point masses, but does not tell us how to model rotation. For rigid bodies, angular velocities proceeds like this: $\boldsymbol{w}_\times \mathbb{I}\boldsymbol{w} + \mathbb{I}\dot{\boldsymbol{w}} = \boldsymbol{\tau}_{ext}$. That's the rotation element of the Newton-Euler equations.

To understand the Newton-Euler equations we will write down the Lagrangian for a rigid body with rotation.

This document aims to provide a from-scratch derivation of the Newton-Euler equations, including the inertia matrix and the gyroscopic term.

It further aims to include many preliminaries, with exceptions noted here.

2 Preliminaries

Table 1: Units and Dimensions

Name	Description	Space	Units
$K_{\mathbf{v}}$	(Linear) kinetic energy	$\mathbb{R}$	$\frac{kg \cdot m^2}{s^2} = J$
K_ω	(Rotational) kinetic energy	$\mathbb{R}$	J
K_{tot}	Total kinetic energy	$\mathbb{R}$	J
$\boldsymbol{v}$	Velocity of a point on the body	$\mathbb{R}^3$	m/s
$\boldsymbol{r}$	Position relative to center of mass	$\mathbb{R}^3$	m
$\dot{\boldsymbol{x}}$	Velocity of center of mass	$\mathbb{R}^3$	m/s
$\boldsymbol{w}$	Angular velocity of body in world frame	$\mathbb{R}^3$	rad/s
ρ	Density of a differential piece of the body at r	$\mathbb{R}$	kg/m^3
$d\Omega$	differential piece of the body	$\mathbb{R}$	m^3
L	Radius of displacement of a point on a body	$\mathbb{R}$	m

2.1 The Cross Product

The cross product of $\mathbf{w}$ and $\mathbf{p}$, $\mathbf{w} \times \mathbf{p}$ is a linear operation which yields a vector $\mathbf{q}$ that is perpendicular to both $\mathbf{w}$ and $\mathbf{p}$.

Because the cross product is a linear operation, we can write it as a matrix. The matrix form makes reasoning quite a bit more convenient. If $\boldsymbol{w} = (w_1, w_2, w_3)$ then

$$\boldsymbol{w}_\times \equiv \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \tag{1}$$

Sometimes $\mathbf{w}_\times$ is written as $[\mathbf{w}]_\times$. In words you'll hear it called $skew(\mathbf{w})$

So $(\mathbf{w}_\times)\mathbf{p} = \mathbf{w} \times \mathbf{p} = cross\_product(\mathbf{w}, \mathbf{p})$

This makes it clear that the cross product has three fun properties

- It's skew-symmetric: $\boldsymbol{w}_\times^T = -\boldsymbol{w}_\times$.
- It is anticommutative: $\boldsymbol{w}_\times \boldsymbol{p} = -\boldsymbol{p}_\times \boldsymbol{w}$
- And it satisfies the Jacobi Identity: $(\boldsymbol{w}_\times \boldsymbol{p})_\times = \boldsymbol{w}_\times \boldsymbol{p}_\times - \boldsymbol{p}_\times \boldsymbol{w}_\times$

It's a convenient trick for computing derivatives of expressions involving the cross product. Example:

$$\frac{d}{d\boldsymbol{v}}(\boldsymbol{w} \times \boldsymbol{v}) = \frac{d}{d\boldsymbol{v}} \boldsymbol{w}_\times \boldsymbol{v} = \boldsymbol{w}_\times \tag{2}$$

2.2 Angular Velocity

Angular velocity is an angle-axis quantity. Consider a body rotating with a constant angular velocity, $\boldsymbol{\omega}$. The unit-vector $\hat{\mathbf{n}} = \boldsymbol{\omega} / \|\boldsymbol{\omega}\|$ is the axis about which the body rotates. $\|\boldsymbol{\omega}\|$ $[rad/s]$ is the rate of rotation about the axis. In other words, $\|\boldsymbol{\omega}\|$ is the angle (in radians) that the body will rotate about the axis $\hat{n}$ in one second.

Given a linear velocity $\mathbf{v}$ you can compute the future position of a point $\mathbf{p}$ as $\mathbf{p}(t) = \mathbf{v}t + \mathbf{p}(0)$.

Constructing a future rotation from an angular velocity is more involved. Consider a rotation. Choice of representation, be it quaternion, matrix, euler angles, rotor or whatever doesn't matter. This rotation transforms points in the body frame to the world frame ${}^{World}\mathbf{R}_{Body}$, read as "R world from body". To keep equations concise, we will use only the first letter of each frame: ${}^{W}\mathbf{R}_{B}$.

We compute a point in the world frame given a point in the body frame by multiplying: $\mathbf{p}_{_W} = {}^{W}\mathbf{R}_{B} * \mathbf{p}_{_B}$

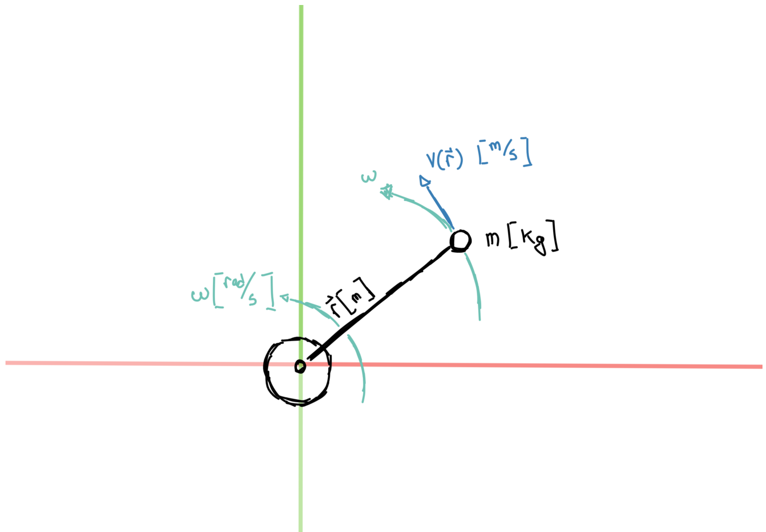
We can compose that rotation with an angular velocity using the 'exp' function. Assume exp converts an angle-axis rotation into whatever representation you happen to be using.

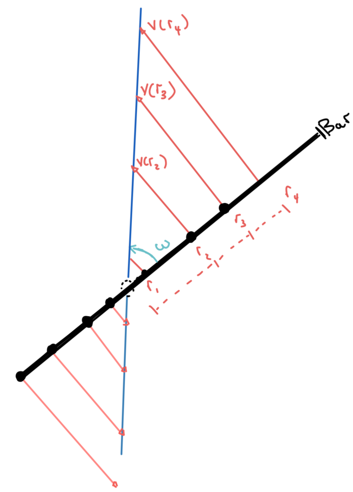
Given an angular velocity in the body frame, $\boldsymbol{\omega}_{_B}$, then ${}^{W}\mathbf{R}_{B}(t) = {}^{W}\mathbf{R}_{B}(0) exp(t * \boldsymbol{\omega}_{_B})$

Similarly, given an angular velocity in the world frame, $\boldsymbol{\omega}_{_W}$, then ${}^{W}\mathbf{R}_{B}(t) = exp(t * \boldsymbol{\omega}_{_W})\, {}^{W}\mathbf{R}_{B}(0)$

2.2.1 Linear Velocity of Points on a Rotating Body

Each point on a rigid body has the same angular velocity. Instantaneously, every point on a rotating body has a linear velocity, coming from the gross motion of the body and its rotation.





Consider a point at a distance $L[m]$ along the x-axis of a rigid bar. As the rigid bar rotates, it makes an angle with the x-axis. We'll call that angle $\theta[rad]$. As the rigid bar translates, the position of all of the points on the bar translate the same amount. We'll track a single point on the rigid body, the center of mass.

Let's compute the time derivative of the position of the point in the world frame.

$$\begin{aligned}
\mathbf{p}(\theta) &= L \cdot \mathbf{r}(\theta) & + \mathbf{x}_{com} \\
\mathbf{p}(\theta) &= L \cdot \begin{bmatrix} sin(\theta) \\ cos(\theta) \end{bmatrix} & + \mathbf{x}_{com} \\
\frac{d}{dt}\mathbf{p}(\theta) = \dot{\mathbf{p}}(\theta) &= L \cdot \begin{bmatrix} -cos(\theta) \\ sin(\theta) \end{bmatrix} \dot{\theta} & + \dot{\mathbf{x}}_{com} \\
\dot{\mathbf{p}}(\theta) &= L \cdot \mathbf{r}(\theta + \frac{\pi}{2})\dot{\theta} & + \dot{\mathbf{x}}_{com}
\end{aligned}$$

Where $\dot{x}_{com}$ is the velocity of the center of mass.

Note the cute trick at the end: We observe that $\dot{r}$ is r rotated clockwise by $\frac{\pi}{2} rads$, scaled by angular velocity.

The three dimensional case can be reduced to two dimensions by Euler's Rotation Theorem. All sequences of rotations in three dimensions can be written as a single rotation about a single axis. This means that, instantaneously, there is a single axis of rotation, which implies that for each point on the body, there is a single plane of rotation, perpendicular to the plane. The velocity in that plane is a rotation of the position vector by $\frac{\pi}{2}$ about the axis of rotation. Conveniently, we can write this as follows:

$$\mathbf{v}(\mathbf{r}) = \dot{\mathbf{p}} = \boldsymbol{\omega} \times \mathbf{r} + \dot{\mathbf{x}}_{com} \tag{3}$$

Sidebar: In my opinion, the cleanest way to derive this in 3D is using the Lie derivative. Ethan Eade's document linked later provides the tools for such a derivation.

2.3 Inertia

In classical mechanics, "inertia" is the second derivative of kinetic energy with respect to velocity for some state. The second derivative matrix is called the "hessian". For both position and orientation, it can be written as a 3x3 matrix.

This is intuitive: inertia tells us, to second order, how much energy is required to change the velocity of some state.

Inertia is also the jacobian of momentum with respect to velocity. Note: Momentum is the gradient of energy with respect to velocity. This is also intuitive: inertia tells us how momentum must change to change velocity.

2.3.1 Positional Inertia

Positional inertia (mass) is the second derivative (hessian) of kinetic energy with respect to linear velocity. The contribution to kinetic energy of linear velocity will be called K_v.

$$
\begin{aligned}
K_v &= \frac{1}{2} m \boldsymbol{v}^T \boldsymbol{v} \\
\frac{\partial}{\partial \boldsymbol{v}} K_v &= m \boldsymbol{v} \\
\frac{\partial^2}{\partial \boldsymbol{v} \partial \boldsymbol{v}} K_v &= \boldsymbol{I} m
\end{aligned}
$$

2.3.2 Rotational Inertia

"Rotational inertia" is the hessian of kinetic energy with respect to angular velocity. We must compute the kinetic energy of a rotating body to see this.

To compute the kinetic energy of a rotating body, we will decompose it into point masses that only have linear velocity. In the model used here, point masses have no defined angular velocity. Then we will sum the kinetic energies of those particles.

If $\dot{x}$ is the velocity of the center of mass of the body, r is a position in the body frame, ω is the angular velocity of the body in the world frame, Ω is the set of positions that comprise the body and ρ is the density for some infinitesmal piece of the body ($d\Omega$) then...

$$
\begin{aligned}
\boldsymbol{v}(\boldsymbol{r}) &= \boldsymbol{w}_\times \boldsymbol{r} + \dot{\boldsymbol{x}} \\
\frac{dK_{tot}}{d\Omega} &= \frac{\rho}{2} \boldsymbol{v}(\boldsymbol{r})^T \boldsymbol{v}(\boldsymbol{r}) \\
K_{total} &= \int_\Omega \frac{\rho(\mathbf{r})}{2} \boldsymbol{v}(\boldsymbol{r})^T \boldsymbol{v}(\boldsymbol{r}) d\Omega \\
K_{tot} &= \int_\Omega \frac{\rho}{2} (\boldsymbol{w}_\times \boldsymbol{r} + \dot{\boldsymbol{x}})^T (\boldsymbol{w}_\times \boldsymbol{r} + \dot{\boldsymbol{x}}) d\Omega \\
K_{tot} &= \int_\Omega \frac{\rho}{2} \left((\boldsymbol{w}_\times \boldsymbol{r})^T (\boldsymbol{w}_\times \boldsymbol{r}) + \dot{\boldsymbol{x}}^T \dot{\boldsymbol{x}} + 2 \dot{\boldsymbol{x}}^T (\boldsymbol{w}_\times \boldsymbol{r}) \right) d\Omega \\
K_{tot} &= \int_\Omega \frac{\rho}{2} (-\boldsymbol{r}_\times \boldsymbol{w})^T (-\boldsymbol{r}_\times \boldsymbol{w}) d\Omega + \int_\Omega \frac{\rho}{2} \dot{\boldsymbol{x}}^T \dot{\boldsymbol{x}} d\Omega + \dot{\boldsymbol{x}}^T \boldsymbol{w}_\times \int_\Omega \rho \boldsymbol{r} d\Omega \\
K_{tot} &= \int_\Omega \frac{\rho}{2} (-\boldsymbol{r}_\times \boldsymbol{w})^T (-\boldsymbol{r}_\times \boldsymbol{w}) d\Omega + \frac{1}{2} m \dot{\boldsymbol{x}}^T \dot{\boldsymbol{x}} + 0 \\
K_{tot} &= \int_\Omega \frac{\rho}{2} (\boldsymbol{r}_\times \boldsymbol{w})^T (\boldsymbol{r}_\times \boldsymbol{w}) d\Omega + K_v \\
K_{tot} &= K_\omega + K_v \\
K_\omega &= \int_\Omega \frac{\rho}{2} (\boldsymbol{r}_\times \boldsymbol{w})^T (\boldsymbol{r}_\times \boldsymbol{w}) d\Omega
\end{aligned}
\tag{4}
$$

The red term is zero, because it is computing the center-of-mass as an offset from the center-of-mass. You can choose different coordinates where the energy cross-term does not disappear. The blue term is linear kinetic energy, the portion of kinetic energy contributed by the linear motion of the whole body.

Now, because angular velocity is constant throughout the body, we can remove it from the integral.

$$\begin{aligned} K_\omega &= \int_\Omega \frac{\rho}{2} \boldsymbol{w}^T \boldsymbol{r}_\times^T \boldsymbol{r}_\times \boldsymbol{w} d\Omega \\ K_\omega &= \frac{1}{2} \boldsymbol{w}^T \left(\int_\Omega \rho \boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega \right) \boldsymbol{w} \\ \mathbb{I} &= \int_\Omega \rho \boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega \\ K_\omega &= \frac{1}{2} \boldsymbol{w}^T \mathbb{I} \boldsymbol{w} \end{aligned} \tag{5}$$

$\mathbb{I}$ is the rotational inertia matrix.

Some intuition: The entries in the rotational inertia matrix will grow larger as $\mathbf{r}$ grows larger. So a spherical shell with all of the mass concentrated on the outside takes more energy to spin than a sphere of uniform density.

The filling of the Three Gorges Dam reservoir moved millions of kilograms of water to a higher altitude than it would otherwise rest. To maintain its kinetic energy, the Earth famously slowed its period of rotation by 0.06 microseconds.

3 Equations of Motion

To reveal the equations of motion, we will apply Hamilton's Principle, the cornerstone of Lagrangian mechanics. While we won't derive it here, Hamilton's Principle can be viewed as a restatement of $\mathbf{F} = m\mathbf{a}$. For every displacement state, there is an associated velocity. For position, it is 'linear velocity', for orientation it is 'angular velocity'. there is a corresponding "generalized force". For position, this generalized force is usually just called "force" and is measured in Newtons. For orientation, the generalized force is called "torque".

"Momentum" for some parameter is the derivative of energy with respect to velocity. "positional" or "linear" momentum is the derivative of energy with respect to velocity. "Angular" momentum is the derivative of energy with respect to angular velocity.

"Generalized force" for some parameter is the derivative of momentum for that parameter with respect to time. For position, that is again, "force", and for rotation, that is torque.

$$\begin{aligned} \frac{d}{dt}\frac{d}{d\dot{\mathbf{x}}} K_{tot} &= m\mathbf{a} = \mathbf{F}_{external} \\ \frac{d}{dt}\frac{d}{d\boldsymbol{w}} K_{tot} &= \boldsymbol{\tau}_{external} \end{aligned} \tag{6}$$

3.0.1 Linear Motion

Linear motion follows the familiar Newton's law.

$$\begin{aligned} \frac{d}{dt}\frac{d}{d\dot{\mathbf{x}}} K_{tot} &= \mathbf{F}_{external} \\ \frac{d}{dt}\frac{d}{d\dot{\mathbf{x}}} K_{tot} &= \frac{d}{dt}\frac{d}{d\dot{\mathbf{x}}} \frac{1}{2} m \dot{\mathbf{x}}^T \dot{\mathbf{x}} \\ &= \frac{d}{dt} m\dot{\mathbf{x}} \\ &= m\ddot{\mathbf{x}} \end{aligned} \tag{7}$$

This is the same as $\mathbf{F} = m\mathbf{a}$

3.0.2 Angular Motion

We will compute the equations of motion by expanding the above definition of torque using expressions we derived earlier.

$$\begin{aligned}
\frac{d}{dt}\frac{d}{d\boldsymbol{\omega}}K_w &= \frac{d}{dt}\frac{d}{d\boldsymbol{\omega}}\frac{1}{2}w^T\mathbb{I}w = \boldsymbol{\tau}_{ext} \\
&= \frac{d}{dt}\mathbb{I}w = \boldsymbol{\tau}_{ext} \\
&= \frac{d}{dt}\left(\int_\Omega \rho \boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega\right) w \\
&= \left(\int_\Omega \rho \frac{d}{dt}\boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega\right) w + \mathbb{I}\dot{w} \\
&= (\int_\Omega \rho(\dot{\boldsymbol{r}}_\times^T \boldsymbol{r}_\times + \boldsymbol{r}_\times^T \dot{\boldsymbol{r}}_\times) d\Omega) w + \mathbb{I}\dot{w} \\
&= \left(\int_\Omega \rho((\boldsymbol{w}_\times \boldsymbol{r} + \dot{\mathbf{x}})_\times^T \boldsymbol{r}_\times + \boldsymbol{r}_\times^T (\boldsymbol{w}_\times \boldsymbol{r} + \dot{\mathbf{x}})_\times)) d\Omega\right) + \mathbb{I}\dot{w} \\
&= \left(\int_\Omega \rho((-\boldsymbol{w}_\times \boldsymbol{r}_\times + \boldsymbol{r}_\times \boldsymbol{w}_\times)\boldsymbol{r}_\times + \dot{\mathbf{x}}_\times^T \mathbf{r}_\times - \boldsymbol{r}_\times(\boldsymbol{w}_\times \boldsymbol{r}_\times - \boldsymbol{r}_\times \boldsymbol{w}_\times) + \mathbf{r}_\times^T \dot{\mathbf{x}}_\times) d\Omega\right) \boldsymbol{w} + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \int_\Omega \rho((-\boldsymbol{w}_\times \boldsymbol{r}_\times + \boldsymbol{r}_\times \boldsymbol{w}_\times)\boldsymbol{r}_\times \boldsymbol{w} - \boldsymbol{r}_\times(\boldsymbol{w}_\times \boldsymbol{r}_\times - \boldsymbol{r}_\times \boldsymbol{w}_\times)\boldsymbol{w} + (\dot{\mathbf{x}}_\times^T \mathbf{r}_\times - \mathbf{r}_\times^T \dot{\mathbf{x}}_\times)\boldsymbol{w}) d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \qquad (8) \\
&= \int_\Omega \rho(-\boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{r}_\times \boldsymbol{w} + \boldsymbol{r}_\times \boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{w} - \boldsymbol{r}_\times \boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{w} - \boldsymbol{r}_\times \boldsymbol{r}_\times \boldsymbol{w}_\times \boldsymbol{w} + (\dot{\mathbf{x}}_\times^T \mathbf{r}_\times - \mathbf{r}_\times^T \dot{\mathbf{x}}_\times)\boldsymbol{w}) d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \int_\Omega \rho(-\boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{r}_\times \boldsymbol{w} + \boldsymbol{r}_\times \boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{w} - \boldsymbol{r}_\times \boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{w} - 0 + (\dot{\mathbf{x}}_\times^T \mathbf{r}_\times - \mathbf{r}_\times^T \dot{\mathbf{x}}_\times)\boldsymbol{w}) d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \int_\Omega \rho(-\boldsymbol{w}_\times \boldsymbol{r}_\times \boldsymbol{r}_\times \boldsymbol{w} + 0) d\Omega + \int_\Omega (\rho(\dot{\mathbf{x}}_\times^T \mathbf{r}_\times - \mathbf{r}_\times^T \dot{\mathbf{x}}_\times)\boldsymbol{w}) d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \int_\Omega \rho(\boldsymbol{w}_\times \boldsymbol{r}_\times^T \boldsymbol{r}_\times \boldsymbol{w}) d\Omega + \int_\Omega \rho(\dot{\mathbf{x}}_\times^T \mathbf{r}_\times \boldsymbol{w} - \mathbf{r}_\times^T \dot{\mathbf{x}}_\times \boldsymbol{w}) d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \boldsymbol{w}_\times \left(\int_\Omega \rho \boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega\right) \boldsymbol{w} + -\dot{\mathbf{x}}_\times^T \boldsymbol{w}_\times \int_\Omega \rho \mathbf{r} d\Omega + (\dot{\mathbf{x}}_\times \boldsymbol{w})_\times \int_\Omega \rho \mathbf{r} d\Omega + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \boldsymbol{w}_\times \left(\int_\Omega \rho \boldsymbol{r}_\times^T \boldsymbol{r}_\times d\Omega\right) \boldsymbol{w} + -0 + 0 + \mathbb{I}\dot{\boldsymbol{w}} \\
&= \boldsymbol{w}_\times \mathbb{I}\boldsymbol{w} + \mathbb{I}\dot{\boldsymbol{w}} = \boldsymbol{\tau}_{ext}
\end{aligned}$$

Where $\boldsymbol{\tau}_{ext}$ is Torque.

The cyan terms are zero because they are (again) computing the center-of-mass as an offset from the center-of-mass, and that offset is zero. The red term is zero because $w_\times w = 0$. We used the Jacobi identity to free expressions of the form $(\boldsymbol{\omega}_\times \mathbf{r} + \dot{\mathbf{x}}_\times^T \mathbf{r})_\times$

The term with the cross product is called "Torque-Free Precession". This is very surprising. Nothing like the cross-product term exists for translation! The wacky cross-product term only appears in three dimensions with bodies that are not radially symmetric. Notice that even if **no** external force is applied, that

$$\dot{\boldsymbol{w}} = \mathbb{I}^{-1}\boldsymbol{w}_\times \mathbb{I}\boldsymbol{w} \qquad (9)$$

With no drag, this means that an angular velocity that is not aligned with one of the axes of inertia will rotate about one of the axes. In practice, damping will gradually cause angular velocity to align with one of the principal axes of inertia.

Torque free precession is the cause of unintuitive behavior like the stability of spinning tops.

For any system where For systems where ω is an eigenvector of $\mathbb{I}$, there will be no torque-free precession.

4 Appendix A

4.1 Properties of the Cross Product Matrix

FAQ: Matrix multiplication is associative, but the cross product is not. How can we create a matrix form of the cross product?

A: The cross-product matrix forces an order of operations.

$$\begin{aligned} a \times (b \times c) &= a_\times b_\times c \\ (a \times b) \times c &= (a_\times b)_\times c \\ &= (a_\times b_\times - b_\times a_\times)c \end{aligned}$$

It is poor form to provide an expression like $a \times b \times c \times d$ without fully parenthesizing each factor.

4.2 Derivatives of Rotations and Functions of Rotations

See http://ethaneade.com/lie.pdf section 2.4.2 for a description of the derivative expression.

4.3 Special Thanks

Special thanks to Ethan Lipson, Drew Bagnell, Jason Nezvadovitz and Ethan Eade for their careful review.

References

New methods for the determination of the motion of rigid bodies (Nova methodus motum corporum rigidorum degerminandi), L. Euler, 1752