

Classical Mechanics

Haydn Cheng

Aug 2024

Contents

1	Newtonian mechanics	2
1.1	Point particle	2
1.2	System of particles	4
1.3	Rigid Body Mechanics	6
1.3.1	Prerequisites	6
1.3.2	Angular velocity vector	7
1.4	Tensor of Inertia	9
1.4.1	Angular Momentum and Energy	9
1.4.2	Parallel Axis Theorem	10
1.4.3	Euler's Equations	11
1.5	Torque-free Precession	12
1.5.1	Euler Angles	16
1.6	Lagrangian mechanics	18
1.6.1	Constraints	18
	Appendices	20
A	Rigid Body Mechanics	21
A.1	Chasles' Theorem	21
A.2	Noncommutability of finite rotations	21

Newtonian mechanics

1.1 Point particle

To solve for the motion of objects, one may employ the vectorial approach with Newton's second law

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{dm\mathbf{v}}{dt}. \quad (1.1)$$

In a special case where the mass of the body is changing, the law becomes

$$\mathbf{F} = m\mathbf{a} - \mathbf{v}_{rel} \frac{dm}{dt} \quad (1.2)$$

which can be derived easily by using eq. (1.1) in an infinitesimal time interval (also known as the force-impulse theorem) as depicted [here](#).

Of course, eq. (1.1) still works but eq. (1.2) provides more useful information that allows us to solve for the motion of the mass. A classic example would be the rocket equation.

To account for the rotational motion of an object, torque $\boldsymbol{\tau}$ and angular momentum \mathbf{L} ¹ are introduced, defined by

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F} \quad \text{and} \quad \mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (1.3)$$

We can see that both $\boldsymbol{\tau}$ and \mathbf{L} depend on the origin defined as \mathbf{r} is the position vector.

¹One may question the necessity to introduce the concept of torque and angular momentum. Indeed, with Newton's second law, one can virtually solve all mechanics problems without resorting to other physical laws. However, when analyzing rigid bodies with spatial extent (in contrast to a point particle), torque becomes useful because the internal forces in these bodies are generally very complicated. In fact, $\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt}$ is merely an extension of Newton's second law as explained and derived [here](#). With a different viewpoint, Noether's theorem dictates that since the universe is rotationally symmetric, so \mathbf{L} must be conserved, in some sense \mathbf{L} is just some useful conserved quantity that is a consequence of a certain symmetry just like how the Laplace-Runge-Lenz vector is the repercussion of some hidden symmetry in higher dimensions.

Taking the derivatives of the angular momentum, we yield “Newton’s second law for rotation”

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \mathbf{r} \times \frac{d\mathbf{p}}{dt} + \frac{d\mathbf{r}}{dt} \times \mathbf{p} = \mathbf{r} \times \mathbf{F} + \mathbf{v} \times (m\mathbf{v}) = \mathbf{r} \times \mathbf{F} = \boldsymbol{\tau} \quad (1.4)$$

Another quantity that is introduced to simplify the matter (which ultimately comes from the symmetry of time) is kinetic energy T and potential energy U defined by

$$T = \frac{1}{2}mv^2 \text{ and } \mathbf{F} = -\nabla V \quad (1.5)$$

This definition is motivated when considering the work done on a constant mass m by a net force \mathbf{F}_{net}

$$W = \int_1^2 \mathbf{F}_{net} \cdot d\mathbf{r} = \int_1^2 \frac{dm\mathbf{v}}{dt} \cdot d\mathbf{r} = m \int_1^2 d\mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = \frac{m}{2} \int_1^2 d(v^2) = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 \quad (1.6)$$

so that we can say the work done by the net force is equal to the change in the kinetic energy (also known as the work-energy theorem)

$$W = W_{con} + W_{non-con} = \Delta T.^2 \quad (1.7)$$

If we define $\mathbf{F}_{net} = \mathbf{F}_{con} + \mathbf{F}_{non-con}$ which is comprised of both conservative forces such as gravity where $\mathbf{F} \propto \frac{\mathbf{r}}{r^2}$ as well as non-conservative forces such as friction.

The defining properties of conservative forces are:

1. $\oint \mathbf{F} \cdot d\mathbf{r} = 0$, or equivalently, from the Stoke’s theorem,
2. $\nabla \times \mathbf{F} = 0$, which both imply
3. The work done by conservative forces is independent of the path taken, as if the work done by the conservative force from point 1 to 2 is a constant value and by switching the sign of $d\mathbf{r}$ in $W = \int_1^2 \mathbf{F} \cdot d\mathbf{r}$, we see that the work done from point 2 to 1 adds a negative sign to that constant value and thus the work done of a loop is zero which is equivalent to $\oint \mathbf{F} \cdot d\mathbf{r} = 0$.

Thus from the vector identity $\nabla \times (\nabla V) = 0$ and the second item above ($\nabla \times \mathbf{F} = 0$), we can define the potential energy as mentioned and the work-energy theorem (eq. (1.7)) becomes the conservation of energy

$$W_{non-con} = T + V. \quad (1.8)$$

²Some authors use ΔW to denote work done, however, as work done should not be interpreted as changes, which would be meaningless, the Δ symbol is omitted. Formally, dW is used to denote the inexact differential, but the complexity of the symbol forbids me to consistently type it in latex.

1.2 System of particles

Having laid out the rudimentary principles, we now investigate the motion of a system of particles.

The translational equation of motion of the i th particle is

$$\mathbf{F}_i = \sum_j \mathbf{F}_{j \rightarrow i} + \mathbf{F}_{i,ext} = \frac{d^2(m_i \mathbf{r}_i)}{dt^2}. \quad (1.9)$$

Summing over all particles,

$$\sum_i \sum_j \mathbf{F}_{j \rightarrow i} + \sum_i \mathbf{F}_{i,ext} = \mathbf{F}_{ext,net} = \sum_i \frac{d^2(m_i \mathbf{r}_i)}{dt^2} = \frac{d^2}{dt^2} \sum_i m_i \mathbf{r}_i = \left(\sum_i m_i \right) \ddot{\mathbf{R}}, \quad (1.10)$$

where we have used Newton's third law, stating that $\mathbf{F}_{i \rightarrow j} = -\mathbf{F}_{j \rightarrow i}$ and

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{\sum_i m_i} \quad (1.11)$$

is defined as the position vector of the center of mass of the system.

This tells us that the total linear momentum of the system is the same as if the entire mass were concentrated at the center of mass and moving with it

Now for the rotational equation of motion of the i th particle, we have

$$\mathbf{r}_i \times \mathbf{F}_i = \mathbf{r}_i \times \mathbf{F}_{i,ext} + \mathbf{r}_i \times \sum_j \mathbf{F}_{j \rightarrow i} = \frac{dL_i}{dt}. \quad (1.12)$$

Summing over all particles,

$$\sum_i (\mathbf{r}_i \times \mathbf{F}_{i,ext}) + \sum_i (\mathbf{r}_i \times \sum_j \mathbf{F}_{j \rightarrow i}) = \boldsymbol{\tau}_{ext} + \sum_i ((\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{j \rightarrow i}) = \boldsymbol{\tau}_{ext} = \sum_i \frac{dL_i}{dt} = \dot{\mathbf{L}}_{tot}. \quad (1.13)$$

where we again used Newton's third law and assumed that the internal forces are central, *i.e.* the force between two particles act on the line connecting them.

To express \mathbf{L}_{tot} in a more convenient form, we define $\mathbf{r}'_i = \mathbf{r}_i - \mathbf{R}$ ³ as shown in fig. 1.1

³We will adopt this convention for the rest of this set of notes

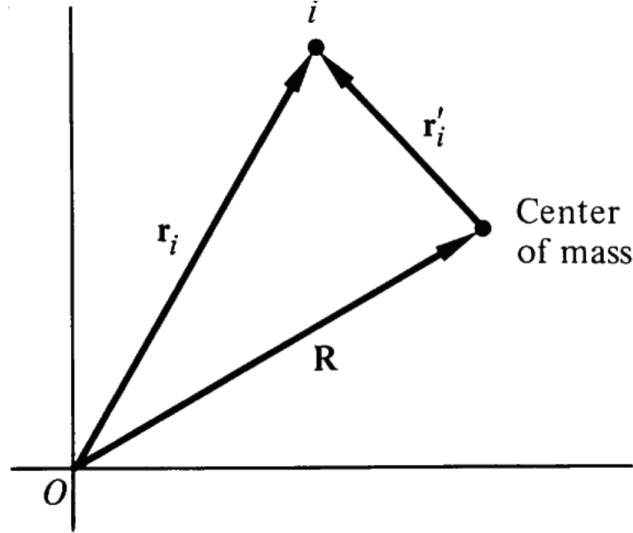


Figure 1.1

, thus $\mathbf{p}_i = m_i \dot{\mathbf{r}}'_i + m_i \dot{\mathbf{R}}$ and the total angular momentum becomes

$$\mathbf{L}_{tot} = \sum_i \mathbf{r}_i \times \mathbf{p}_i = \sum_i ((\mathbf{r}'_i + \mathbf{R}) \times (m_i \dot{\mathbf{r}}'_i + m_i \dot{\mathbf{R}})) = (\sum_i m_i) \mathbf{R} \times \dot{\mathbf{R}} + \sum_i m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) \quad (1.14)$$

where the cross terms $\sum_i (\mathbf{R} \times m_i \dot{\mathbf{r}}'_i + \mathbf{r}'_i \times m_i \dot{\mathbf{R}})$ are omitted since $\sum_i m_i \mathbf{r}'_i = 0$ from the definition of the center of mass.

So we see that the total angular momentum of a system of particles (due to $\dot{\mathbf{r}}$) can be split into two parts. The first term is due to the orbital motion of the center of mass about the origin due to transnational motion (due to $\dot{\mathbf{R}}$) and the second is due to the spinning motion of the particles around their center of mass (due to $\dot{\mathbf{r}}'$).

The same reasoning applies to the kinetic energy for a system of particles, where one term is attributed to the collective movement, while another arises from the rotational motion about the center of mass

$$T = \sum_i \frac{1}{2} m_i v_i^2 = \frac{1}{2} \sum_i m_i (\dot{\mathbf{r}}'_i + \dot{\mathbf{R}})^2 = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}'^2_i + \frac{1}{2} (\sum_i m_i) \dot{\mathbf{R}}^2 \quad (1.15)$$

where we neglect the cross term $\sum_i m_i (\dot{\mathbf{r}}'_i \cdot \dot{\mathbf{R}})$ for the same reason explained above.

virial
theorem
var

1.3 Rigid Body Mechanics

1.3.1 Prerequisites

If one were to choose a theorem that represents the crux of rigid body motion, one would have to pick Chasles' Theorem, which states that it is always possible to describe an arbitrary displacement of a rigid body by a translation of its center of mass plus a rotation around its center of mass (it can rotate about an arbitrary point but the center of mass is the most convenient choice).^{4 5} The formal proof requires complex matrix algebra but a simple way to demonstrate the theorem is given in section A.1. Since the translational and the rotational motion of a rigid body are separable, so we almost always assume that the translational motion has already been accounted for. In fact, we will assume that the center of mass is at rest for the rest of this section.

If one were to pick a second theorem, then it would be Euler's Theorem, which states that any displacement of a rigid body such that a point on the rigid body remains fixed is equivalent to a single rotation about some axis that runs through the fixed point. Since the center of mass is always fixed as established above, it tells us that rotation about the center of mass means that all points on the rigid body undergo circular motion with respect to the closest point on an axis that runs through the center of mass where the direction of the axis defines the rotational motion and is in the same direction as the angular velocity which will be explored more in the next subsection. The proof of Euler's Theorem will not be given here due to its complexity.

Before diving into the physics of rigid body motion, some conventions of notations used in this set of notes should be explicitly stated first, as different texts would use different notations.

1. The spaced-fixed coordinate system, which is stationary in the lab frame has axes ($\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$) which obeys the right-hand rule. Quantities observed from the lab frame (or the space frame) are the same as quantities measured from the spaced-fixed coordinate system.
2. The body-fixed coordinate system has axes ($\hat{\mathbf{1}}$, $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$) which also follows the right-hand rule and always coincide with the principle axes of the body. Quantities observed from the body frame are the same as quantities measured from the body-fixed coordinate system.
3. The instantaneous inertial frame with axes labeled $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ is an inertial frame which its axes coincide with the body's principal axes only at time instant t . This frame is not rotating with respect to the lab frame so it is equally superior.
4. The Euler angles that are used to transform between these two coordinate systems

⁴Another way to construct any displacement is first to do a rotation and then translate parallel to the axis of rotation, we reverse the order of translation and rotation while adding a constraint on the translating direction.

⁵Yet another interesting and useful fact is that if the motion of the body is planar (*i.e.* the angular velocity is perpendicular to the linear velocity), then there always exists an instantaneous axis of rotation (which need not be inside the rigid body) that is parallel to the angular velocity such that any infinitesimal displacement can be constructed by rotating around this axis (This is the third way in which a displacement of a rigid body can be constructed). The proof of this fact is given [here](#). The instantaneous axis of rotation can be constructed geometrically mentioned in idea 33 of [this handout](#) by Jaan Kalda.

are rotated in the $z - x - z$ sequence.

1.3.2 Angular velocity vector

Before handling the rather complicated mathematical treatments, it is useful to define what do we mean by angular velocity.

Angular velocity, similar to linear velocity, is a quantity describing a body's (more rigorously, the body-fixed frame's) motion that is independent of the choice of a coordinate system or origin. One may imagine there is an “angular-speedometer” that can measure the angular velocity of a rigid body undergoing any general motion. However, it is frame-dependent, meaning that the angular velocity observed in the lab frame is different from that observed from another.

Suppose we have 3 orthogonal frames: the lab frame, which is not rotating and fixed in space ⁶. And two other frames whose origins remained fixed (as our interests do not lie on the translational motions and rotational motions can be analyzed separately from translational motions) and can rotate freely about their origins. Each of these two frames possess their own angular velocity vector as observed from the fixed lab frame, which passes through their origins and the direction defines their rotational motion as guaranteed by Euler's theorem, where every points co-rotating with the frame trace out a circle with the center at the closet point to the rotation axis.

As linear velocity is defined as the time derivative of the displacement vector, one may be tempted to define an “angular displacement vector”, describing how an object undergo rotation and the angular velocity can be simply defined as the time derivative of the “angular displacement vector”. However, this is not possible for the fact that finite rotations do not commute in 3-dimensional space (for 2D case, rotations do commute as there are only 2 degrees of freedom which can be assigned to positive and negative signs) as one can play with literally any object to try it out, so

$$\Delta\boldsymbol{\theta} \stackrel{?}{=} \Delta\theta_x\hat{\mathbf{x}} + \Delta\theta_y\hat{\mathbf{y}} \neq \Delta\theta_y\hat{\mathbf{y}} + \Delta\theta_x\hat{\mathbf{x}}. \quad (1.16)$$

However, we *can* define an “infinitesimal angular displacement vector” as angular infinitesimal displacements do commute (less obvious but one still gets a feeling by playing with an object but limiting the angles rotated to be very small), so

$$\delta\boldsymbol{\theta} = \delta\theta_x\hat{\mathbf{x}} + \delta\theta_y\hat{\mathbf{y}} = \delta\theta_y\hat{\mathbf{y}} + \delta\theta_x\hat{\mathbf{x}}. \quad (1.17)$$

To prove the above result, we consider fig. 1.2. Without loss of generality, we define the z -axis of the lab frame (which is arbitrarily defined) to coincide with the angular velocity vector of the rotating frame, and the \mathbf{r} vector to be the position vector of any point co-rotating with the rotating frame. The direction of rotation $\delta\boldsymbol{\theta}$ can be x or y axis in the above equation.

⁶From the similarity between angular velocity and linear velocity, one may think there is no universally superior frame of reference when analyzing rotational motion due to relativity. However, rotation is absolute as one may determine whether it is rotating from local measurement, e.g. whether the equipotential surface of a bucket of water is parabolic or horizontal. Although there is still debate on this topic, e.g. [here](#), we take this fact for granted as we are still in the realm of Newtonian physics.

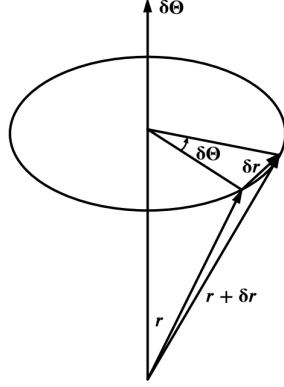


Figure 1.2

From it, it is clear that

$$\delta \mathbf{r} = \delta \boldsymbol{\theta} \times \mathbf{r}. \quad (1.18)$$

Considering two successive rotation through $\delta \boldsymbol{\theta}_1$ and $\delta \boldsymbol{\theta}_2$, we have

$$\delta \mathbf{r}_{12} = \delta \boldsymbol{\theta}_1 \times \mathbf{r} + \delta \boldsymbol{\theta}_2 \times (\mathbf{r} + \delta \mathbf{r}) = (\delta \boldsymbol{\theta}_1 + \delta \boldsymbol{\theta}_2) \times \mathbf{r} = \delta \mathbf{r}_{21} \quad (1.19)$$

if we neglect the higher-order term. An alternate proof providing more intuition but more tedious is given in section A.2.

Dividing eq. (1.18) by δt , we have

$$\mathbf{v} = \frac{\delta \mathbf{r}}{\delta t} = \frac{\delta \boldsymbol{\theta}}{\delta t} \times \mathbf{r}. \quad (1.20)$$

In a more general case where the origin is moving at a velocity \mathbf{v}_O , then the velocity of point P in the rigid body will be

$$\mathbf{v}_P = \mathbf{v}_O + \boldsymbol{\omega} \times \mathbf{r}_{O \rightarrow P} \quad (1.21)$$

which is a very useful equation since it relates the velocity of any two points in the rigid body with the common angular velocity (note that O need not be the center of mass of the rigid body, as if true, $\mathbf{v}_P = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m. \rightarrow P}$ and $\mathbf{v}_O = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m. \rightarrow O}$ yields the general result. In fact, eq. (1.21) can be regarded as the mathematical definition for the angular velocity vector.

Another very useful property of the angular velocity vector is that the law of angular velocity addition to find the relative angular velocity between different frames is exactly analogous to the law of linear velocity addition, where

$$\boldsymbol{\Omega}_{1\text{rel.}2} = \boldsymbol{\Omega}_{1\text{rel.}3} - \boldsymbol{\Omega}_{2\text{rel.}3}. \quad (1.22)$$

To prove this, we first define clearly what do we mean by relative velocity in the linear case. Suppose we have a point P_1 co-rotating with S_1 and P_2 fixed in S_2 . In the lab frame S_3 , the displacement vectors of P_1 and P_2 are defined as the changes in their linear positions as measured in the lab frame. The linear velocity vectors are defined as the displacement vectors divided by a regular time interval, and the relative velocity of the 2 points (or 2 frames) is the difference in their linear velocity vectors. In the angular velocity case, we can simply follow the same argument as “angular displacement vector” is well defined as long as the time interval concerned tends to zero. It is helpful to visualize the “angular displacement vector” in the 2D case, where the time interval concerned is not limited to infinitesimally small, then it becomes clear that angular velocity vectors do add like linear velocity vectors by considering the most simple case: S_1 rotating with the angular velocity $\omega_1 \hat{\mathbf{z}}$ and S_2 with $\omega_2 \hat{\mathbf{z}}$, then after a time interval Δt , the angular displacement vectors are $\boldsymbol{\theta}_1 = \omega_1 \Delta t \hat{\mathbf{z}}$ and $\boldsymbol{\theta}_2 = \omega_2 \Delta t \hat{\mathbf{z}}$ and the relative displacement vector is $\boldsymbol{\theta}_{1rel,2} = (\omega_1 - \omega_2) \Delta t \hat{\mathbf{z}}$ thus the relative angular velocity vector is $\boldsymbol{\omega}_{1rel,2} = (\omega_1 - \omega_2) \hat{\mathbf{z}}$. The same thing applies in our 3D world, just that $\Delta \rightarrow \delta$ and it is harder to visualize the entire process.

1.4 Tensor of Inertia

1.4.1 Angular Momentum and Energy

Now we return to eq. (1.14) and try to evaluate the abstract summation form of the spin angular momentum due to rotation about the center of mass $\mathbf{L}_{rot} = \sum_i m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i)$ when the rigid body is rotating about its center of mass at an angular velocity $\boldsymbol{\omega}$.

Now from fig. 1.2 we can conclude the general relationship that if a vector \mathbf{r} is rotating about a fixed origin with angular velocity $\boldsymbol{\omega}$, then we have the relation

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r}. \quad (1.23)$$

Therefore, \mathbf{L}_{rot} becomes

$$\mathbf{L}_{rot} = \sum_i m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) = \sum_i m_i (\mathbf{r}'_i \times (\boldsymbol{\omega} \times \mathbf{r}'_i)) = \sum_i m_i (|\mathbf{r}'_i|^2 \boldsymbol{\omega} - \mathbf{r}'_i (\mathbf{r}'_i \cdot \boldsymbol{\omega})). \quad (1.24)$$

From here, we can explicitly write out the x, y and z components of \mathbf{L}_{rot} as

$$\begin{aligned} L_{rot,x} &= \sum_i m_i ((x_i'^2 + y_i'^2 + z_i'^2) \omega_x - x_i' (x_i' \omega_x + y_i' \omega_y + z_i' \omega_z)) \\ &= \sum_i m_i ((y_i'^2 + z_i'^2) \omega_x - (x_i' y_i') \omega_y - (x_i' z_i') \omega_z), \\ L_{rot,y} &= \sum_i m_i (x_i' (x_i'^2 + z_i'^2) \omega_y - (y_i' z_i') \omega_z - (x_i' y_i') \omega_x), \\ \text{and } L_{rot,z} &= \sum_i m_i ((x_i'^2 + y_i'^2) \omega_z - (x_i' z_i') \omega_x - (y_i' z_i') \omega_y). \end{aligned} \quad (1.25)$$

In matrix form,

$$\mathbf{L}_{rot} = \begin{pmatrix} L_{rot,x} \\ L_{rot,y} \\ L_{rot,z} \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \tilde{\mathbf{I}}\boldsymbol{\omega}. \quad (1.26)$$

Similarly, for the abstract sum for the kinetic energy in eq. (1.15) due to the rotational motion, it now becomes

$$T_{rot} = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i'^2 = \frac{1}{2} \sum_i m_i (\boldsymbol{\omega} \times \mathbf{r}_i') \cdot (\boldsymbol{\omega} \times \mathbf{r}_i') = \frac{1}{2} \boldsymbol{\omega} \cdot \sum_i m_i \mathbf{r}_i' \times (\boldsymbol{\omega} \times \mathbf{r}_i') = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} \quad (1.27)$$

where we used the vector identity $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$.

Now one of the great advantages of the use of principal axes is the simplification of eq. (1.27), as it now becomes

$$T_{rot} = \frac{1}{2} I_{xx} \omega_x^2 + \frac{1}{2} I_{yy} \omega_y^2 + \frac{1}{2} I_{zz} \omega_z^2. \quad (1.28)$$

var 13.7
diagonal-
ization
and iner-
tia ten-
sor

1.4.2 Parallel Axis Theorem

If the tensor of inertia about the center of mass $\tilde{\mathbf{I}}_{c.m.}$ and the displacement vector pointing from the center of mass to point $P = X\hat{\mathbf{i}} + Y\hat{\mathbf{j}} + Z\hat{\mathbf{k}}$ are known, then the tensor of inertia about point P will be

$$\begin{aligned} I_{xx,P} &= \sum_i m_i (y_{i,P}'^2 + z_{i,P}'^2) = \sum_i m_i ((y'_{i,c.m.} - Y)^2 + (z'_{i,c.m.} - Z)^2) \\ &= \sum_i m_i ((y_{i,c.m.}'^2 + z_{i,c.m.}'^2) + (Y^2 + Z^2) - 2(y_{i,c.m.}Y + z_{i,c.m.}Z)) \\ &= I_{xx,c.m.} + \sum_i m_i (Y^2 + Z^2) \end{aligned} \quad (1.29)$$

$$\begin{aligned} \text{and } I_{xy,P} &= - \sum_i m_i (x_{i,P} y_{i,P}') = \sum_i m_i ((x_{i,c.m.} - X)(y_{i,c.m.} - Y)) \\ &= \sum_i m_i ((x_{i,c.m.} y_{i,c.m.}) - XY - (X y_{i,c.m.} + x_{i,c.m.} Y)) \\ &= I_{xy,c.m.} - \sum_i m_i XY \text{ etc.} \end{aligned}$$

where the last term in each of the equations vanishes due to the property of the center of mass.

1.4.3 Euler's Equations

With all the prerequisites explained, we are now ready to tackle the seemingly simple differential equation $\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt}$. Consider a time instant t when a rigid body is rotating with $\boldsymbol{\omega}$. Since the body frame is non-inertial, we cannot apply this rotational Newton's law here. What we can do, however, is to consider an inertial frame that only coincides with the body frame at time t .

It is very important to have this picture in mind: at time t , the inertial frame axes $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ are the same as the body axes $\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}$. Then, after time dt , the axes of the body-fixed coordinate system rotate by an angle of $\boldsymbol{\omega}dt$ along $\boldsymbol{\omega}$ while the inertial frame axes remained stationary. So from the inertial frame, the body axes actually rotate with $\boldsymbol{\omega}$. We then repeat this procedure infinite time.

Writing out the equation of motion in this inertial coordinate system, we have

$$\begin{aligned}\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} &= \frac{d}{dt}(L_1\hat{\mathbf{1}} + L_2\hat{\mathbf{2}} + L_3\hat{\mathbf{3}}) = \frac{dL_1}{dt}\hat{\mathbf{1}} + L_1\frac{d\hat{\mathbf{1}}}{dt} + \frac{dL_2}{dt}\hat{\mathbf{2}} + L_2\frac{d\hat{\mathbf{2}}}{dt} + \frac{dL_3}{dt}\hat{\mathbf{3}} + L_3\frac{d\hat{\mathbf{3}}}{dt} \\ &= \frac{dL_1}{dt}\hat{\mathbf{1}} + \frac{dL_2}{dt}\hat{\mathbf{2}} + \frac{dL_3}{dt}\hat{\mathbf{3}} + (\boldsymbol{\omega} \times \hat{\mathbf{1}})L_1 + (\boldsymbol{\omega} \times \hat{\mathbf{2}})L_2 + (\boldsymbol{\omega} \times \hat{\mathbf{3}})L_3\end{aligned}\tag{1.30}$$

where $i = 1, 2$ and 3 and we used eq. (1.23) since the body-fixed axes ($\hat{\mathbf{1}}, \hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$) are rotating angular velocity $\boldsymbol{\omega}$ about the inertial instantaneous frame as mentioned.

Splitting the vector equation into three components, we have three non-linear coupled first-order differential equations

$$\begin{aligned}\tau_1 &= I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) \\ \tau_2 &= I_2\dot{\omega}_2 + \omega_3\omega_1(I_1 - I_3) \\ \tau_3 &= I_3\dot{\omega}_3 + \omega_1\omega_2(I_2 - I_1).\end{aligned}\tag{1.31}$$

An alternate derivation of Euler's equations with discrete time interval considerations can be found in Chapter 8.7.2 of Kleppner. An alternate proof of Euler's equations by the Euler-Lagrange equation can be found in Chapter 13.18 of Cline.

One has to be reminded that although the set of equations are given in body-fixed coordinates and thus are only valid at time t where the body frame coincides with the inertial frame, since t is arbitrarily chosen, the equations of motion tell us things that are more general than the behaviors of the system at that mere instant. In fact, we can create an infinite number of instantaneous inertial frames such that Euler's equations are always valid. In retrospect, the introduction of an instantaneous inertial frame was merely to derive Euler's equations and nothing more. From now on there are only 2 frames that matter: the body frame and the lab frame.

Also, since the Euler equations only depend on the principal moments of inertia I_1, I_2 and I_3 , thus all bodies having the same principal moments of inertia will behave exactly the same even though the bodies may have very different shapes. The simplest geometrical shape of a body having three different principal moments is a homogeneous ellipsoid. Thus, the

rigid body motion often is described in terms of the equivalent ellipsoid that has the same principal moments of inertia.

Example: Kleppner (3rd. ed) Example 8.16

Question: Due to $\boldsymbol{\omega}$ not necessarily parallel to \mathbf{L} , many peculiar phenomena are observed in rigid body motion. One of which is the Tennis Racket Theorem (also known as the Intermediate axis theorem), which states that the rotations about the 2 principal axes which have the largest and the smallest moment of inertia are stable while the rotation about the intermediate axis is not. Prove it.

Solution:

To explain this phenomenon, we suppose that the body initially spins with $\boldsymbol{\omega} = \omega_1 \hat{\mathbf{e}}_1$ and receives small perturbations on ω_2 and ω_3 . Then according to the Euler's equations, we have $\omega_1 = \text{constant}$ and

$$\frac{d^2\omega_2}{dt^2} + \left(\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3} \omega_1^2 \right) \omega_2 = 0 \quad (1.32)$$

as one can easily verify. So we see that ω_2 undergo simple harmonic motion if I_1 is the largest or the smallest moment of inertia, but increase exponentially with time and the motion is unstable.^a

^aFor a more intuitive explanation, refer to the [explanation](#) given by the famous mathematician Terrance Tao as well as this [video](#) by the famous YouTuber Veritasium.

Example: Kleppner (3rd. ed) Example 8.17

Question: A uniform rod is mounted on a horizontal frictionless axle through its center. The axle is carried on a turntable rotating at a constant angular velocity $\boldsymbol{\Omega}$ as depicted in fig. 1.3 . Find $\theta(t)$ shown in the figure.

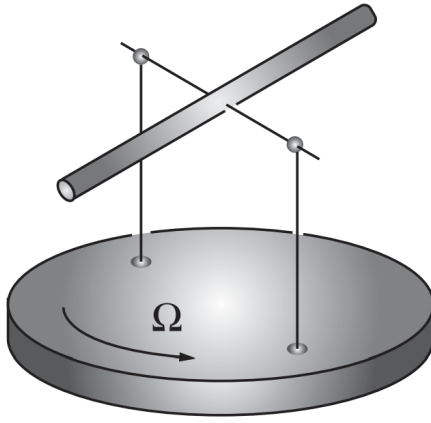
Solution: Referring to the figures, we have $\omega_1 = \dot{\theta}$, $\omega_2 = \Omega \sin \theta$ and $\omega_3 = \Omega \cos \theta$. Substituting them into the Euler's equations and leveraging the small angle approximation $\sin \theta \approx \theta$ gives

$$\ddot{\theta} + \left(\frac{I_3 - I_2}{I_1} \right) \Omega^2 \theta = 0. \quad (1.33)$$

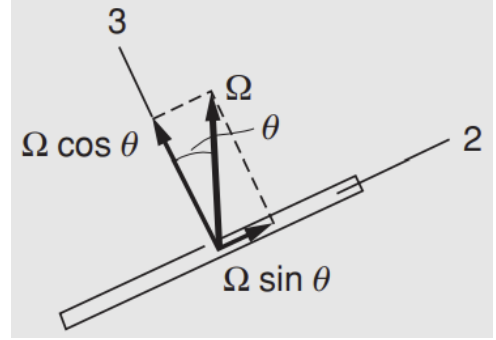
So we conclude that θ undergo simple harmonic motion with angular frequency $\gamma = \sqrt{\frac{I_3 - I_2}{I_1}} \Omega$.

1.5 Torque-free Precession

One of the most classic applications of Euler's equations is a torque-free precession. Consider a symmetric top with I_1 being the moment of inertia about the symmetric axis and $I_2 = I_3 = I_\perp$. Then the equations give $\omega_1 = \text{constant} = \omega_s$ and



(a)



(b)

Figure 1.3

$$\frac{d^2\omega_2}{dt^2} + \left(\frac{I_1 - I_\perp}{I_\perp}\right)^2 \omega_s^2 \omega_2 = 0. \quad (1.34)$$

So ω_2 undergo simple harmonic motion with angular frequency $\gamma = \left| \frac{I_1 - I_\perp}{I_\perp} \right| \omega_s$

$$\omega_2 = \omega_\perp \cos \gamma t \quad (1.35)$$

where ω_\perp depends on the initial condition .

Further calculation would give that

$$\omega_3 = \pm \omega_\perp \sin \gamma t \quad (1.36)$$

where the positive sign corresponds to the case where $I_1 > I_\perp$ indicates the body is short and fat so the spin is clockwise, and vice versa.

To get qualitatively what really happens, refer to fig. 1.4 .

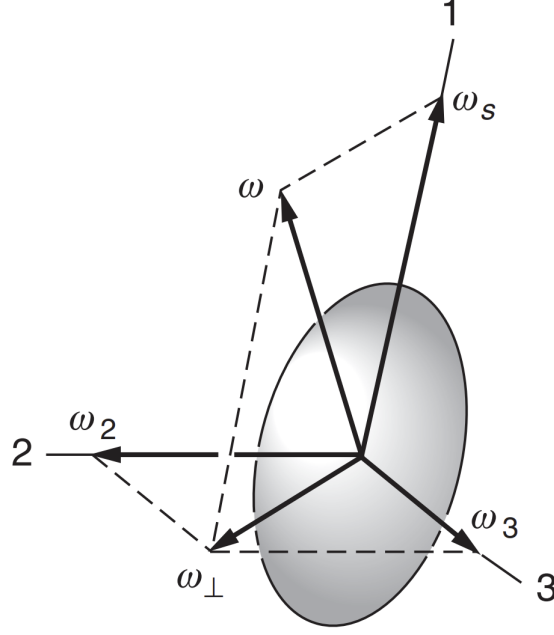


Figure 1.4

$\omega_1 = \omega_s = \text{constant}$ simply means that in the body frame, the component of $\boldsymbol{\omega}$ on $\hat{\mathbf{1}}$ has a fixed magnitude ω_s .⁷

The solution for ω_2 and ω_3 means that they are actually components of $\boldsymbol{\omega}_\perp$ on $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ respectively when $\boldsymbol{\omega}_\perp$ is rotating about $\hat{\mathbf{1}}$ at the angular speed γ when observed in the body frame.

Combining these two insights, we can say that $\boldsymbol{\omega}_\perp$ rotate about $\hat{\mathbf{1}}$ at the angular speed $\gamma + \omega_s$ when observed from the lab frame by simple angular velocity addition.

Furthermore, since $I_2 = I_3 = I_\perp$ and $\mathbf{L}_2 = I_2\boldsymbol{\omega}_2$ and $\mathbf{L}_3 = I_3\boldsymbol{\omega}_3$ therefore $\mathbf{L}_\perp = \mathbf{L}_2 + \mathbf{L}_3 = I_\perp(\boldsymbol{\omega}_2 + \boldsymbol{\omega}_3) = I_\perp\boldsymbol{\omega}_\perp$ which means that $\hat{\mathbf{1}}, \boldsymbol{\omega}_1 = \omega_s\hat{\mathbf{1}}, \mathbf{L}_1 = I_1\boldsymbol{\omega}_1, \boldsymbol{\omega}_\perp, \mathbf{L}_\perp = I_\perp\boldsymbol{\omega}_\perp, \boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_\perp, \mathbf{L} = \mathbf{L}_1 + \mathbf{L}_\perp$ are all in the same plane, and since $\hat{\mathbf{1}}$ is fixed in the body frame, the only degree of freedom is that all the vectors mentioned above rotate about $\hat{\mathbf{1}}$ with the same angular speed.⁸ But we already found out that one of the vectors, namely $\boldsymbol{\omega}_\perp$ has an angular speed of γ , so all the vectors mentioned have the same angular velocity $\gamma\hat{\mathbf{1}}$.

We have already solved the problem in the body frame, next we transform it back into the lab frame, which is what we care about the most.

In the space-fixed inertial frame, since there are no external torques in torque-free

⁷This also means that at every time instant t , $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ revolve about $\hat{\mathbf{e}}_1$ (technically not $\hat{\mathbf{1}}$ since $\hat{\mathbf{1}}$ is not fixed in the instantaneous inertial frame so it is meaningless to talk about rotation around $\hat{\mathbf{1}}$ in this frame and also $\hat{\mathbf{1}}, \hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are relatively fixed so no axis is rotating about another axis but since $\hat{\mathbf{1}}$ and $\hat{\mathbf{e}}_1$ coincide at that moment, this saying is generally accepted) at constant angular speed ω_s when observed from the instantaneous inertial frame.

⁸The fact that the angles between all the vectors are fixed is trivial in the body frame considering the mathematical form of each vector listed above. To prove this fact in space frame, notice that \mathbf{L} of the body is fixed in *torque-free* precession, and we have shown that $\omega_1 = \omega_s = \text{constant}$ and $\boldsymbol{\omega}_\perp$ is constant as well, so α shown in fig. 1.5 must be constant. To be extra cautious, we can say since $T_{rot} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{L} = \frac{1}{2}\omega L \cos \alpha$ (eq. (1.27)) must be constant since there is no external work done, so α must be constant.

precession, \mathbf{L} is now fixed in place.

From the analysis done in the body frame, we must bear this fact in mind: all the vectors concerned in this problem are in the same plane. To visualize, it is helpful to imagine that all the vectors are on a piece of paper with $\hat{\mathbf{l}}$ and $\boldsymbol{\omega}_\perp$ being the two adjacent edges of the paper and \mathbf{L} being the diagonal (it is always possible since the size of the paper is arbitrary). In the body frame, $\hat{\mathbf{l}}$ is held still so the piece of paper rotates about one vertical edge with angular speed γ similar to how a door rotates about a door hinge.

However, refer to fig. 1.5 where now we wish to fix \mathbf{L} in place in space frame meaning that the 2 corners (the tip and the tail of \mathbf{L}) are now stationary and the piece of paper rotates about \mathbf{L} . This picture explains intuitively why although $\boldsymbol{\omega}$ and \mathbf{L} has the same angular velocity in the body frame but when switched to the lab frame, where \mathbf{L} is fixed, $\boldsymbol{\omega}$ is not fixed but is now co-rotating with $\hat{\mathbf{l}}$ about \mathbf{L} with the same angular speed Ω_p . Mathematically, \mathbf{L} is also rotating with the angular speed Ω_p , just that the axis of rotation is \mathbf{L} itself, so it is equivalent to having no rotation at all.

To find this new common angular speed Ω_p , we can utilize the angular velocity addition formula eq. (1.22), where frame 1 is a frame where $\boldsymbol{\omega}$ is at rest, frame 2 is lab frame and frame 3 is the body frame. So

$$\boldsymbol{\Omega}_{\boldsymbol{\omega} \text{ rel. lab}} = \boldsymbol{\Omega}_{\boldsymbol{\omega} \text{ rel. body}} - \boldsymbol{\Omega}_{\text{lab rel. body}}. \quad (1.37)$$

or

$$\Omega_p \hat{\mathbf{z}} = \gamma \hat{\mathbf{l}} - (-\boldsymbol{\omega}). \quad (1.38)$$

Resolving this vector equation along $\hat{\mathbf{l}}$ gives

$$\begin{aligned} \Omega_p \cos \alpha &= \gamma + \omega_s \\ \Omega_p &= \frac{I_1 \omega_s}{I_\perp \cos \alpha}. \end{aligned} \quad (1.39)$$

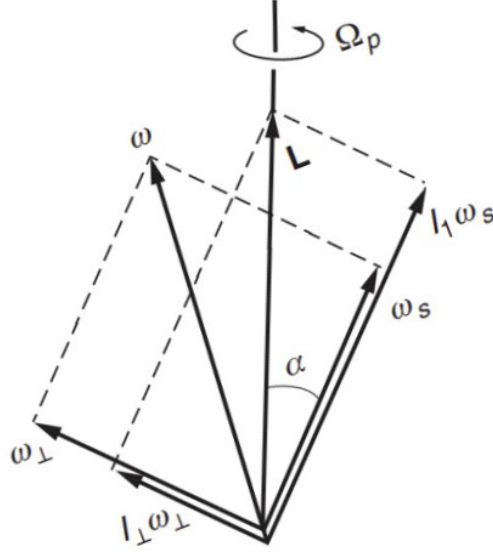


Figure 1.5

The intuitive explanation as to why $\Omega_p \cos \alpha = \gamma + \omega_s$ is as follows:

Firstly, as mentioned, $\omega_1 = \omega_s = \text{constant}$ implies that $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ revolve around $\hat{\mathbf{1}}$ (technically, $\hat{\mathbf{e}}_1$) at ω_s . However, even then, we have calculated that $\boldsymbol{\omega}_\perp$ (and also $\boldsymbol{\omega}$ and other relevant vectors) still have angular speed γ in the body frame where $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are at rest. This means that those sets of vectors rotate at the angular speed $\gamma + \omega_s$ about $\hat{\mathbf{1}}$ in the lab frame.

Secondly, we resort to the “2D paper model” developed above. We now know that for the “door hinge” mode (rotate about $\hat{\mathbf{1}}$), the angular speed observed from the lab frame is $\gamma + \omega_s$. We want to know what the angular speed observed from the lab frame is when rotating about \mathbf{L} . To answer this, we have to remember the vector property of angular velocity $\boldsymbol{\Omega}_p$. We utilize this fact and resolve $\boldsymbol{\Omega}_p$ along $\hat{\mathbf{1}}$ (and its perpendicular direction). The former angular speed which equals to $\Omega_p \cos \alpha$ should be identical to the angular speed of the set of vectors when $\hat{\mathbf{1}}$ is fixed which we calculated to be $\gamma + \omega_s$.

1.5.1 Euler Angles

The description of rigid body rotation is greatly facilitated by transforming from the space-fixed (lab) coordinates $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ to the body-fixed coordinates $(\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}})$ since the inertia tensor measured with this coordinate is always diagonal. They can be related, as introduced in the “Maths”, by

$$(\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}) = \boldsymbol{\lambda}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}). \quad (1.40)$$

As mentioned in “Maths”, only 3 independent angles are needed for any rotational transformation. By convention, the Euler angles ϕ, θ, ψ are used. Refer to fig. 1.6.

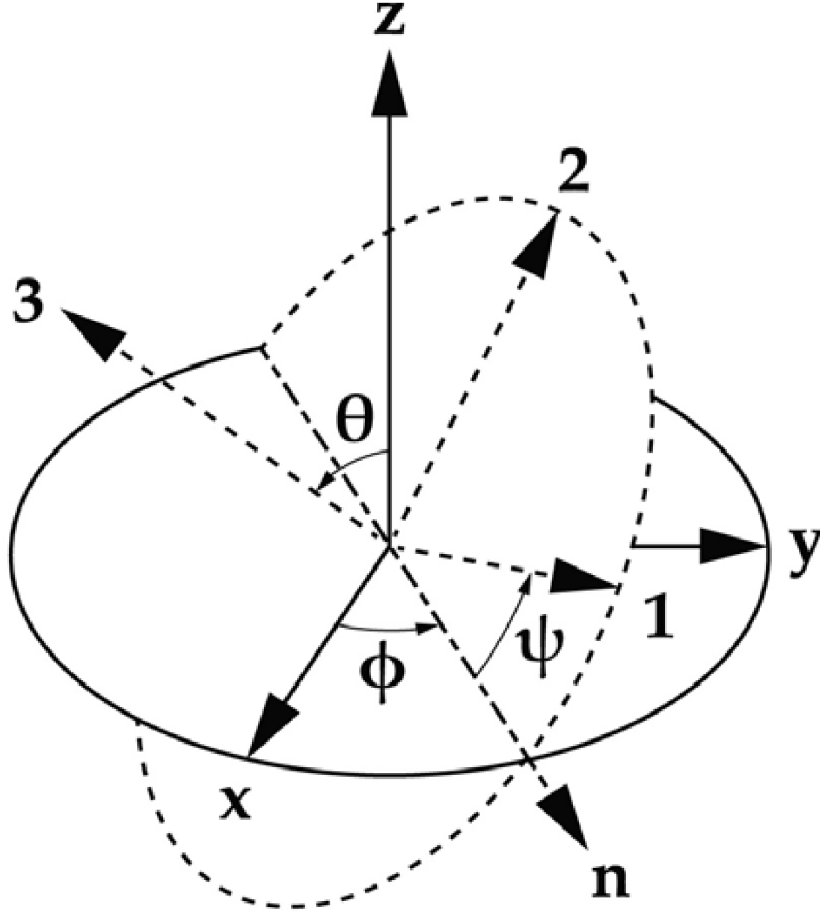


Figure 1.6

The unit vector defined by $\hat{\mathbf{n}} = \hat{\mathbf{z}} \times \hat{\mathbf{3}}$ is called the line of nodes.

Firstly, $\hat{\mathbf{x}}$ is made to coincide with the line of node $\hat{\mathbf{n}}$, then while keeping $\hat{\mathbf{x}}$ unchanged, $\hat{\mathbf{z}}$ is made to coincide with $\hat{\mathbf{3}}$ (which is possible since the line of node is defined to be $\hat{\mathbf{n}} = \hat{\mathbf{z}} \times \hat{\mathbf{3}}$). Lastly, while keeping $\hat{\mathbf{z}}$ unchanged, $\hat{\mathbf{x}}$ is made to be coincide with 1 axis. As $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are in place, due to the orthogonality of the systems, $\hat{\mathbf{y}}$ is bound to coincide with $\hat{\mathbf{2}}$.

The rotational matrices of each rotation are

$$\boldsymbol{\lambda}_\phi = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\lambda}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1.41)$$

Therefore the total rotational matrix is

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_\phi \boldsymbol{\lambda}_\theta \boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos \phi \cos \psi - \sin \phi \cos \theta \sin \psi & \sin \phi \cos \psi + \cos \phi \cos \theta \sin \psi & \sin \theta \sin \psi \\ -\cos \phi \sin \psi - \sin \phi \cos \theta \cos \psi & -\sin \phi \sin \psi + \cos \phi \cos \theta \cos \psi & \sin \theta \cos \psi \\ \sin \phi \sin \theta & -\cos \phi \sin \theta & \cos \theta \end{pmatrix}. \quad (1.42)$$

The angular velocity will be

$$\boldsymbol{\omega} = \dot{\phi}\hat{\mathbf{z}} + \dot{\theta}\hat{\mathbf{n}} + \dot{\psi}\hat{\mathbf{3}}. \quad (1.43)$$

Expressing $\hat{\mathbf{z}}$ and $\hat{\mathbf{n}}$ in terms of the body-fixed coordinates, we have

$$\begin{aligned} \hat{\mathbf{z}} &= \sin \theta \sin \psi \hat{\mathbf{1}} + \sin \theta \cos \psi \hat{\mathbf{2}} + \cos \theta \hat{\mathbf{3}} \\ \hat{\mathbf{n}} &= \cos \psi \hat{\mathbf{1}} - \sin \psi \hat{\mathbf{2}}. \end{aligned} \quad (1.44)$$

So

$$\boldsymbol{\omega} = (\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \hat{\mathbf{1}} + (\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \hat{\mathbf{2}} + (\dot{\psi} + \dot{\phi} \cos \theta) \hat{\mathbf{3}}. \quad (1.45)$$

By playing a similar game, the angular velocity can be expressed in terms of the space-fixed coordinates, with

$$\boldsymbol{\omega} = (\dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi) \hat{\mathbf{x}} + (\dot{\theta} \sin \phi - \dot{\psi} \sin \theta \cos \phi) \hat{\mathbf{y}} + (\dot{\phi} + \dot{\psi} \cos \theta) \hat{\mathbf{z}}. \quad (1.46)$$

The validity of the results can be verified by confirming that the dot product of $\boldsymbol{\omega}$ with itself

$$\boldsymbol{\omega} \cdot \boldsymbol{\omega} = \omega_1^2 + \omega_2^2 + \omega_3^2 = \omega_x^2 + \omega_y^2 + \omega_z^2 = \dot{\phi}^2 + \dot{\theta}^2 + \dot{\psi}^2 + 2\dot{\phi}\dot{\psi} \cos \theta \quad (1.47)$$

is an invariant under coordinates transformation as any scalar properties like mass, Lagrangian, or Hamiltonian should.

The advantage of working in the body-fixed coordinates is that the inertia tensor is diagonal, which greatly simplifies the work needed in expressing the kinetic energy as

$$T_{rot} = \frac{1}{2}(I_1(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)^2 + I_2(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)^2 + I_3(\dot{\phi} \cos \theta + \dot{\psi})^2). \quad (1.48)$$

1.6 Lagrangian mechanics

1.6.1 Constraints

If the conditions of constraint can be expressed as equations in the form

$$f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, t) = 0, \quad (1.49)$$

then the constraints are said to be holonomic. For example, the constraint for a rigid body is that the distance of any two points in the body is fixed, *i.e.*

$$|\mathbf{r}_i - \mathbf{r}_j| = c_{ij}. \quad (1.50)$$

With Cartesian coordinates, there are initially $3N$ degrees of freedom thus $3N$ independent equations. If there are k holonomic constraints, then we can use them to eliminate variables and solve for the $3N$ variables as well as the k constraint forces. In another viewpoint, we can introduce $3N - k$ independent variables (also known as generalized coordinates ⁹) $q_1, q_2, \dots, q_{3N-k}$ in terms of the old coordinates $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ and time.

In contrary, a constraint on the particle's motion that cannot be expressed in the above form is called a nonholonomic constraint. For example, consider a disk with radius a rolling on the horizontal $x - y$ plane such that the plane of the disk is always vertical. In addition to the x and y coordinates, we use θ to specify the angle between the axis of the disk and the x axis and ϕ to denote the angle of rotation about the axis of the disk. The constraint in this case is that the relative velocity of the contact point between the disk and the plane must be zero, *i.e.*

$$\begin{aligned}
(\dot{x}\hat{\mathbf{x}} + \dot{y}\hat{\mathbf{y}}) + \dot{\phi}(\cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}}) \times (-a\hat{\mathbf{z}}) &= 0 \\
(\dot{x}\hat{\mathbf{x}} + \dot{y}\hat{\mathbf{y}}) + a\dot{\phi}\cos\theta\hat{\mathbf{y}} - a\dot{\phi}\sin\theta\hat{\mathbf{x}} &= 0 \\
(\mathbf{x} - a\dot{\phi}\sin\theta)\hat{\mathbf{x}} + (\mathbf{y} + a\dot{\phi}\cos\theta)\hat{\mathbf{y}} &= 0 \\
\mathbf{x} = a\dot{\phi}\sin\theta \quad \text{and} \quad \mathbf{y} = -a\dot{\phi}\cos\theta.
\end{aligned} \tag{1.51}$$

However, neither of these equations can be integrated without first solving the problem itself, so there are nonholonomic.

⁹In some sense, generalized coordinates are just the minimum numbers of variables which can satisfy the whole system.

Appendices

Rigid Body Mechanics

A.1 Chasles' Theorem

We begin by considering two masses m_1 and m_2 located at \mathbf{r}_1 and \mathbf{r}_2 respectively connected by a thin, rigid and massless rod.

The “rigid body condition” is that the distance between the two masses remained unchanged, *i.e.*

$$\begin{aligned}
 d(|\mathbf{r}_1 - \mathbf{r}_2|) &= 0 \\
 |\mathbf{r}_1 - \mathbf{r}_2| &= c \\
 |\mathbf{r}_1 - \mathbf{r}_2|^2 &= (\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = c^2 \\
 d((\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2)) &= 2(\mathbf{r}_1 - \mathbf{r}_2) \cdot d(\mathbf{r}_1 - \mathbf{r}_2) = 0 \\
 d\mathbf{r}_1 &= d\mathbf{r}_2 \text{ or } (d\mathbf{r}_1 - d\mathbf{r}_2) \perp (\mathbf{r}_1 - \mathbf{r}_2)
 \end{aligned} \tag{A.1}$$

Now since $d\mathbf{r}'_1 = d\mathbf{r}_1 - d\mathbf{R} = (\frac{m_2}{m_1+m_2})(d\mathbf{r}_1 - d\mathbf{r}_2)$ and $d\mathbf{r}'_2 = d\mathbf{r}_2 - d\mathbf{R} = -(\frac{m_1}{m_1+m_2})(d\mathbf{r}_1 - d\mathbf{r}_2)$, so when $d\mathbf{r}_1 = d\mathbf{r}_2$ in the first case, it means that the body undergo pure translation without rotating. And the second case corresponds to a case of translation plus rotation since

1. $d\mathbf{r}'_1 \perp (\mathbf{r}_1 - \mathbf{r}_2)$:

$$d\mathbf{r}'_1 \cdot (\mathbf{r}_1 - \mathbf{r}_2) = (\frac{m_2}{m_1+m_2})(d\mathbf{r}_1 - d\mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0 \tag{A.2}$$

2. $d\mathbf{r}'_2 \perp (\mathbf{r}_1 - \mathbf{r}_2)$: the proof is the same as above

3. $\frac{d\mathbf{r}'_1}{r'_1} = -\frac{d\mathbf{r}'_2}{r'_2}$:

$$\frac{d\mathbf{r}'_1}{r'_1} = (\frac{m_2}{m_1+m_2}) \frac{(d\mathbf{r}_1 - d\mathbf{r}_2)}{r'_1} = (\frac{m_1}{m_1+m_2}) \frac{(d\mathbf{r}_1 - d\mathbf{r}_2)}{r'_2} = -\frac{d\mathbf{r}'_2}{r'_2}. \tag{A.3}$$

A.2 Noncommutability of finite rotations

fig. A.1 illustrates the essence of the general proof of this fact, where we consider the rotation of the position vector $\mathbf{r} = r\hat{\mathbf{i}}$ through an angle α about the z axis and β about the

y axis but in different order. Rotating about z axis by an angle α , $\hat{\mathbf{i}}$ becomes $\cos \alpha \hat{\mathbf{i}} + \sin \alpha \hat{\mathbf{j}}$ while rotating about y axis by an angle β , $\hat{\mathbf{i}}$ becomes $\cos \beta \hat{\mathbf{i}} - \sin \beta \hat{\mathbf{k}}$, so

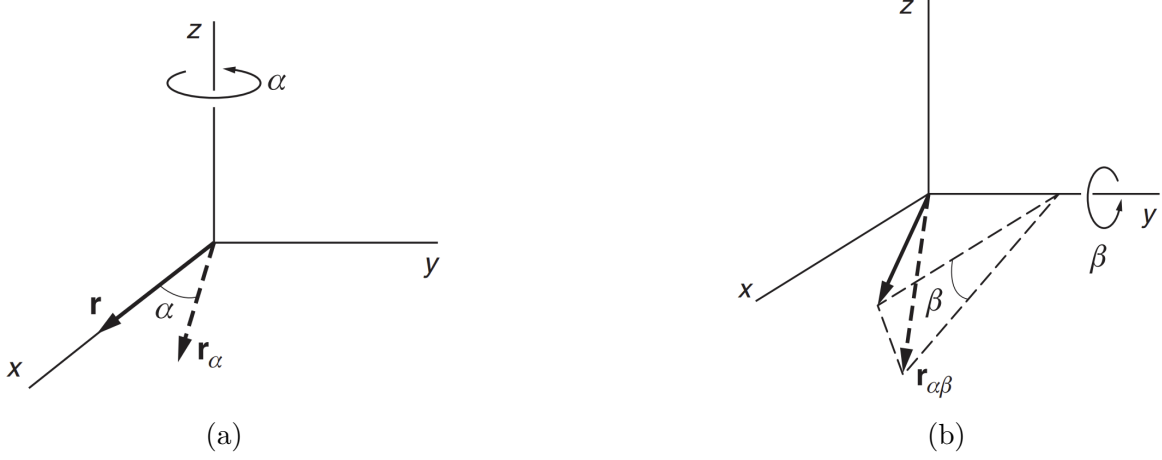


Figure A.1

$$\begin{aligned} \mathbf{r}_{\alpha\beta} &= r \cos \alpha (\cos \beta \hat{\mathbf{i}} - \sin \beta \hat{\mathbf{k}}) + r \sin \alpha \hat{\mathbf{j}} = r \cos \alpha \cos \beta \hat{\mathbf{i}} + r \sin \alpha \hat{\mathbf{j}} - r \cos \alpha \sin \beta \hat{\mathbf{k}} \\ \text{and } \mathbf{r}_{\beta\alpha} &= r \cos \alpha \cos \beta \hat{\mathbf{i}} + r \cos \beta \sin \alpha \hat{\mathbf{j}} - r \sin \beta \hat{\mathbf{k}}. \end{aligned} \quad (\text{A.4})$$

It is evident that while finite size of α and β would result in a difference between $\mathbf{r}_{\alpha\beta}$ and $\mathbf{r}_{\beta\alpha}$, but if we take the limit $\alpha \ll 1$ and $\beta \ll 1$, then $\mathbf{r}_{\alpha\beta} = \mathbf{r}_{\beta\alpha}$ and the angular displacement vector $\Delta \boldsymbol{\theta} = \Delta \alpha \hat{\mathbf{k}} + \Delta \beta \hat{\mathbf{j}}$ is well defined. In particular, the displacement of \mathbf{r} is

$$\Delta \mathbf{r} = \mathbf{r}_{\alpha\beta} - \mathbf{r} = \mathbf{r}_{\beta\alpha} - \mathbf{r} = r \alpha \hat{\mathbf{j}} - r \beta \hat{\mathbf{k}} = \Delta \boldsymbol{\theta} \times \mathbf{r}. \quad (\text{A.5})$$

The linear velocity will then be

$$\mathbf{v} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{r}}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \boldsymbol{\theta} \times \mathbf{r}}{\Delta t} = \boldsymbol{\omega} \times \mathbf{r}. \quad (\text{A.6})$$

where the angular velocity vector $\boldsymbol{\omega}$ is defined as

$$\boldsymbol{\omega} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \boldsymbol{\theta}}{\Delta t} \quad (\text{A.7})$$

In this case, $\boldsymbol{\omega} = \frac{d\beta}{dt} \hat{\mathbf{j}} + \frac{d\alpha}{dt} \hat{\mathbf{k}}$.